DOCTOR OF MEDICINE

Clinical Drug-Gene and Drug-Drug-Gene Interactions for the Most Commonly Used Chronic Drugs in the UK

Malki, Mustafa Adnan

Award date:
2021

Clinical Drug-Gene and Drug-Drug-Gene Interactions for the Most Commonly Used Chronic Drugs in the UK

By Mustafa Adnan Malki

Doctor of Philosophy (PhD) Degree in Medicine

Major Research Area: Clinical Pharmacy

Specific Research Area: Clinical Pharmacogenomics

University of Dundee, UK

2021

Table of Contents

Contents

Chapter I (literature review): Drug-Gene/Drug-Drug-Gene Interactions and Adverse Drug Reactions

Chapter II:
General Methodology

Chapter III:
Drug-Gene Interactions for the Most Commonly Used Chronic Drugs in the UK

Chapter IV: Drug-Drug-Gene Interactions for the Most Commonly Used Chronic Drug Combinations in the UK

Chapter V: Replication Results from the UKBB Primary Care Data for the top Drug-Gene and Drug-Drug-Gene Interactions

Chapter VI:
Final General Discission

Tables and Figures

Chapter I

Tables

Figures

Chapter II

Tables

Figures

Chapter III

Tables

Figures

Chapter IV

Tables

Chapter V

Tables

Figures

Chapter VI

Tables

Figures

This page was left intentionally

Abbreviations

No.	Abbreviation	Meaning
1	ABCB1	ATP Binding Cassette Subfamily B Member 1
2	ABCC1	ATP Binding Cassette Subfamily C Member 1
3	ABCC2	ATP Binding Cassette Subfamily C Member 2
4	ABCC3	ATP Binding Cassette Subfamily C Member 3
5	ABCC4	ATP Binding Cassette Subfamily C Member 4
6	ABCG2	ATP Binding Cassette Subfamily G Member 2
7	ACEIs	Angiotensin-Converting Enzymes Inhibitors
8	ACSs	Acute Coronary Syndromes
9	ACT	Artemisinin-based Combination Therapy
10	ADRs	Adverse Drug Reactions
11	AERS	Adverse Event Reporting System
12	ARBs	Angiotensin Receptor Blockers
13	AUC	Area Under the Curve
14	BBB	Blood-Brain Barrier
15	BCRP	Breast Cancer Resistant Protein
16	CCBs	Calcium Channel Blockers
17	CES1	Carboxylesterase 1
18	CI	Confidence Interval
19	CNS	Central Nervous System
20	CPIC	The Clinical Pharmacogenetics Implementation Consortium
21	CPsI/III	Coproporphyrins I/III
22	CSO	Chief Scientists Office
23	CVDs	Cardiovascular Diseases
24	CYP	Cytochrome P450
25	DDGIs	Drug-Drug-Gene Interactions
26	DDIs	Drug-Drug Interactions
27	DDMEGIs	Drug-Drug-Metabolizing Enzyme Gene Interactions
28	DDTGIs	Drug-Drug-Transporters' Genes' Interactions
29	DDTGIs	Drug-Drug-Transporters' Genes' Interactions
30	DGIs	Drug-Gene Interactions
31	DHPs	Dihydropyridines
32	DM	Diabetes Mellitus
33	DMARDs	Disease-Modifying Antirheumatic Drugs
34	DMEGIs	Drug-Metabolizing Enzyme Gene Interactions
35	DTGIs	Drug-Transporters' Genes' Interactions
36	ED	Enrichment Design
37	EHRs	Electronic Health Records
38	EOD	Exposure-Only Design
39	FDA	Food and Drug Administration
40	GoDARTs	Genetics of Diabetes Audit and Research in Tayside Scotland
41	GOF	Gain-of-Function
42	GoSHARE	Genetics of Scottish Health Research Register

No.	Abbreviation	Meaning
43	GP	General Practitioner
44	GS: SFHS	Generation Scotland: Scottish Family Health study
45	GWAS	Genome-Wide Association Study
46	HWE	Hardy-Weinberg equilibrium
47	LD	Linkage Disequilibrium
48	LDL	Low-Density Lipoprotein
49	LOF	Loss-of-Function
50	MAF	Minor Allele Frequency
51	MATE1	Multidrug and Toxic Compound Extrusion Protein 1
52	MATE2	Multidrug and Toxic Compound Extrusion Protein 2
53	MEs	Metabolizing Enzymes
54	MHRA	Medicines and Healthcare Products Regulatory Agency
55	MI	Myocardial Infraction
56	MRP1	Multidrug Resistance-Associated Protein 1
57	MRP2	Multidrug Resistance-Associated Protein 2
58	MRP3	Multidrug Resistance-Associated Protein 3
59	NHS	National Health Service
60	NLCs	Nocturnal Leg Cramps
61	NRS	NHS Research Scotland
62	NSAIDs	Non-Steroidal Anti-Inflammatory Drugs
63	OATP1B1 (SLCO1B1)	Organic Anion-Transporting Polypeptide 1B1 (Solute Carrier Organic Anion Transporter Family Member 1B1)
64	OATP1B3 (SLCO1B3)	Organic Anion-Transporting Polypeptide 1B3 (Solute Carrier Organic Anion Transporter Family Member 1B3)
65	OATP2B1 (SLCO2B1)	Organic Anion-Transporting Polypeptide 2B1 (Solute Carrier Organic Anion Transporter Family Member 2B1)
66	OCT1 (SLC22A1)	Organic Cation Transporter 1 (Solute Carrier Family 22 Member 1)
67	OCT2 (SLC22A2)	Organic Cation Transporter 2 (Solute Carrier Family 22 Member 2)
68	OCT3 (SLC22A3)	Organic Cation Transporter 3 (Solute Carrier Family 22 Member 3)
69	OR	Odds Ratio
70	PBPK	Physiologically Based Pharmacokinetic
71	PO	Peripheral Oedema
72	P-gp	P-glycoprotein 1
73	Pharmacodynamic	PD
74	PharmGKB	The Pharmacogenomics Knowledge Base
75	PK	Pharmacokinetic
76	PPIs	Proton Pump Inhibitors
77	PTGS1	Prostaglandin-Endoperoxide Synthase 1
78	SBP	Systolic Blood Pressure
79	SHARE	Scottish Health Research Register
80	SNP	Single Nucleotide Polymorphism
81	TCA	Tricyclic Antidepressant
82	Trs	Transporters
83	UKBB	UK Biobank
84	VIF	Variance Inflation Factor
85	WHO	World Health Organization

Acknowledgements

First of all, I would like to offer countless thanks to my supervisor Professor Ewan Pearson for his incredible effort with me from the first day of this PhD project. He has been so supportive, encouraging, and responsive. Without his valuable and continuous support and guidance throughout this project, I couldn't have achieved this final version of my thesis the way it looks now. I started this project with only the UK Biobank cross-sectional data and GoDARTs cohorts. He has thankfully greatly contributed to making the other two cohorts (GS + GoSHARE) also available to me around one year later, which has definitely added more strength to our work. He has then added more power to my work by making the UK Biobank primary care data directly available to the other research team members and me since it had been released in September 2019. Without this, we couldn't have a replication chapter for our results. He has also guided me to develop the online applications to view our findings which has also added more beauty to our work.

I cannot forget the generous financial support I have received from the Government of Saudi Arabia during my entire scholarship period in the UK. For around six years starting from the English course, passing through the Master degree, and ending with this PhD degree, I have been fully funded academically and medically with having all my daily living expenses covered as well.

I'm really grateful to my parents Adnan Malki and Salwa Azhar who have been so encouraging and supportive especially during my writing-up stage during a hard time of COVID-19 pandemic. Countless thanks to their continuous prayers and patience during my absence while I was in the UK. I also would like to thank all of my seven brothers and sisters for their encouragement. In the top of that, endless thanks to my God who has made all that I have mentioned above real.

Declaration

I declare that I am the only one who has produced the entire work in this PhD thesis. I have undertaken all analyses in this project myself for all of the cohorts used starting from cleaning the data, including extracting the phenotype and genotype information (except for the genotype data from the UK Biobank cohort which was thankfully extracted for me by Dr. Kaixin Zhou), and undertaking the different statistical tests. I also declare that I have developed the two online applications to show our results. I declare that I have authored the entire thesis work, including both the written and the visual content.

I declare that this work has not been submitted before for any other academic degrees in the University of Dundee or any other academic institution anywhere. This work is solely submitted for the PhD degree from the University of Dundee.

Signature:

Mustafa Adnan Malki

Thesis Summary

Objectives: In the present project, I attempted to uncover novel and clinically important drug-gene interactions (DGIs) and drug-drug-gene interactions (DDGIs) among 50 commonly used chronic drugs and 50 commonly used chronic drug combinations in the UK.

Methods: Using the UK Biobank (cross-sectional) cohort and 3 other Scottish cohorts (longitudinal), I have studied the association of 162 genetic variants in important genes with three drug response phenotypes for the 50 selected drugs/combinations. This has generated a total of 48,600 findings divided equally between the two studies (DGIs and DDGIs), which I have made accessible via two online databases. I then undertook further replication for our top findings utilizing the UK Biobank primary care data.

Results: We identify 8 novel associations after Bonferroni correction, 3 of which are replicated or validated in the UK biobank or have other supporting results: The C-allele at rs4918758 in CYP2C9 was associated with a 25% (15-44%) lower odds of dose reduction of quinine, p=1.6×10-5; the A-allele at rs9895420 in ABCC3 was associated with a 46% (24-62%) reduction in odds of dose reduction with doxazosin, p=1.2×10-4, and altered blood pressure response in the UK Biobank; the *CYP2D6*2* variant was associated with a 30% (18 %- 40%) reduction in odds of stopping ramipril treatment, p=1.01×10-5, with similar results seen for enalapril and lisinopril and with other CYP2D6 variants.

I have also detected two other novel findings with directionally consistent results in the replication cohort with p-values close to significance levels (amlodipine-rs868853 (ABCC4)-lower odds for daily dose reduction and clopidogrel-rs12353214 (PTGS1)-decreased drug stopping risk)).

In addition, out of 3 novel DDGIs, one association was validated using an alternative phenotype in UK Biobank. In the discovery cohort, carrying the G allele at rs9516519 (T>G) variant in ABCC4 transporter was linked with a 4.72 (2.44-9.13) times increased risk of stopping bisoprolol or atorvastatin treatments when they were used concomitantly ($p=1.48 \times 10^{-5}$). In the replication cohort, this drug combination was associated with a great SBP reduction (~ 8 mmHg drop in mean SBP ($p < 2 \times 10^{-16}$)) and the presence of the rs9516519 (T>G) variant increased this effect.

Finally, 19 DG associations were identified that replicated previous study findings including but not limited to the association of *CYP2C9**3 with increased gliclazide side effects and the association of *CYP2C8**3 with reduced pioglitazone efficacy. We also report some other novel and potentially important associations from both the DG and DDG interaction studies.

Conclusion: The work in this thesis highlights the value of using large population datasets for pharmacogenomic discovery and has identified novel findings that may impact on clinical care.

This page was left intentionally blank

Chapter I (literature review):

Drug-Gene/Drug-Drug-Gene Interactions and Adverse Drug Reactions

The work related to drug-drug-gene interactions in this chapter has been published

(Malki M, Pearson E. Drug–drug–gene interactions and adverse drug reactions. The Pharmacogenomics Journal. 2019;20(3):355-366, *see supplementary material 1*)

Abstract

The economic and health burden caused by adverse drug reactions has increased dramatically in the last few years. This is likely to be mediated by increasing polypharmacy, which increases the likelihood of drug-drug interactions. Tools utilized by health care practitioners to flag potential adverse drug reactions secondary to drug-drug interactions ignore individual genetic variation, which has the potential to markedly alter the severity of these interactions. To date, there have been limited published studies on the impact of genetic variation on drug interactions. In this introduction, I establish a detailed classification for pharmacokinetic drug-gene /drug-drug-gene interactions and give examples from the literature that support this approach. The increasing availability of real-world drug outcome data linked to genetic bioresources is likely to enable the discovery of previously unrecognized, clinically important drug-gene/drug-drug-gene interactions.

1. Introduction

It was previously and alarmingly reported that adverse drug reactions (ADRs) represent the fourth leading cause of death in the USA [1]. A recent review (2015) showed that 3.6% of patients were admitted to hospitals in Europe due to ADRs, and 10% of patients developed side effects during their in-patient stay [2]. The latest report issued by MiDatabank in cooperation with the Medicines and Healthcare Products Regulatory Agency (MHRA), shows an increasing trend in the number of reported ADRs in the period between 2011 and 2016 across the UK [3]. It has also been estimated that ADRs alone cost the NHS £770M annually [4]. Non-steroidal anti-inflammatory drugs (NSAIDs), diuretics, anticoagulants, and antiplatelets have been recognized to be the major culprits, with prescribing errors being major contributors to medication-related adverse events [5]. The chance of prescribing errors increases when patients undergo multiple treatments; a situation that is highly prevalent in elderly patients [6]. There are a number of factors that influence the occurrence of ADRs secondary to drug interactions, such as age, renal function and other comorbidities. In addition, genetic variation is likely to play a crucial role in the development of ADRs. In his publication "*Inborn errors of metabolism*" (1909/1923), the British scientist Archibald Garrod was the first to describe the individual differences between people in metabolizing different substances [7]. This concept had then received more attention in 1956 when some patients with pseudocholinesterase deficiency died after succinylcholine injections [8]. In the following year (1957), Motulsky (USA) has re-defined the concept of Garrod that defects in metabolism may be responsible for variability in drug response [9]. In 1959, the term "*pharmacogenetics*" was first suggested by Friedrich Vogel (Germany) to describe this concept [10].

This term has been then replaced by a broader term *"pharmacogenomics"* to describe the influence of genetic variants in any gene, rather than only metabolizing enzymes genes, to change drug responses between individuals.

The importance of considering genetic variants when evaluating drug outcomes can be observed from one study in which the authors noted that when only considering genetic polymorphisms in three drug-metabolizing enzymes (cytochrome P450 2C9 (CYP2C9), CYP2C19, and CYP2D6), 15% of the ADRs were due to drug-gene interactions (DGIs), and 19% were due to drug-drug-gene interactions (DDGIs) [11]. Incorporation of these gene variants increased the number of predicted clinically critical drug interactions by ~51% [11]. Given the large number of genes involved in drug metabolism and transport, we cannot underestimate the importance of genetic variation in contributing to the potential for clinically critical ADRs.

Following the recent advances in pharmacogenomics, the traditional view of drug-drug interactions needs to be modified to include genetic variation. To date, the literature on drug-gene interactions is still limited although it has been growing rapidly recently while the literature on drug-drug-gene interactions remains very limited, with only one previous review evaluating the impact of *CYP2C9*, *C19* and *2D6* variants [12]. In this introduction, we attempt to provide an in-depth framework for the classification of pharmacokinetic DG/DDG interactions caused by different mechanisms, and their potential impact on increasing clinically critical drug interactions in the context of the polypharmacy seen in modern medicine today.

2. Drug-gene/drug-drug-gene interactions classification

Drug interactions can be divided into three main categories: inhibitory interactions, induction interactions, and phenoconversion interactions. Inhibitory and induction interactions can be defined as any interactions that affect the victim drug's pharmacokinetics (PK) to increase or reduce concentrations of the drug, respectively. Induction or inhibition can occur either with the administration of a perpetrator drug that alters the victim drug metabolism or transport (drug-drug interactions (DDIs)), or with the presence of loss- or gain-of-function (LOF or GOF) genetic variants that alter the function of enzymes that alter metabolism or transport of the victim drug (DGI), or the combination of both (DDGI). A DDGI can be thought of as a double hit – whereby the genetic variant and the perpetrator drug combine to act on transporter or metabolism pathways to greatly alter drug concentrations. It is also possible to see phenoconversion - where the interacting drug effect and the genotype have opposing effects, resulting in a temporary phenotype shift, e.g., neutralizing/reversing the effect of a gain-of-function genotype when an inhibitory drug is prescribed. In this introduction I describe, with examples, different cases of interactions under each of the above three categories, focusing initially on metabolizing enzymes, before considering drug transporters covering examples from both DG and DDG interaction studies.

3. Drug-Metabolizing Enzyme Gene Interactions (DMEGIs)

3.1 Inhibitory Interactions

One of the well-established pharmacogenomic associations is the correlation between warfarin (a CYP2C9 substrate) reduced dosage requirements and carriage of the *CYP2C9*2/*3* reduced activity variants. Carriers of these variant alleles are at increased risk for warfarin-induced bleeding due to decreased warfarin metabolism. In 2007, the Food and Drug Administration (FDA) issued a label informing practitioners about this genotype-phenotype correlation which was updated later in 2010 by providing a specific dosage ranges recommendation [13,14]. In addition, a couple of genotype-guided warfarin dosage estimation tools (http://warfarindosing.org/Source/Home.aspx and http://www.warfarindoserevision.com/the-dose-revision-tool) have been developed to help health care providers to select the most suitable dose based on individual patients' parameters. However, findings from clinical trials regarding the clinical utility of this genotype-phenotype association have been contradictory [15,16].

Another key example for the clinical application of pharmacogenomics is the use of *CYP2D6* genotype data to examine suitability for tamoxifen therapy. Tamoxifen is a prodrug activated by CYP2D6 into its active metabolite endoxifen. The Clinical Pharmacogenetics Implementation Consortium (CPIC) recommends that individuals carrying the CYP2D6 poor metabolizer phenotypes avoid tamoxifen and use alternative breast cancer therapies due to the strong evidence showing lack of efficacy as indicated by the increased risk of breast cancer relapse [17]. In this DG inhibitory interaction, loss of CYP2D6 enzyme function could have resulted in inhibiting the activation of tamoxifen into its active metabolite endoxifen resulting in reducing the therapeutic efficacy of the treatment.

There are different mechanisms by which loss of enzyme function can occur at the molecular (genetic) level. In the coding region of the gene, nonsynonymous single-nucleotide polymorphisms (SNPs) result in changing the sequence of amino acids to produce a malfunctioning protein. This change could occur as a result of replacement of one nucleotide in the codon by another one resulting in a different codon coding for a different amino acid (i.e., missense mutations). Alternatively, nonsynonymous SNPs could also result in producing stop codons, early during the translation process, which prevents the translation of the remaining codons into their corresponding amino proteins (i.e., non-sense mutations). In the non-coding region of the gene, SNPs could affect regions, such as promotors and enhancers, which control transcription processes or could affect non-coding RNAs which could affect gene expression as well.

Poor CYP2D6 metabolizers could represent those who carry one or more variants linked with loss of CYP2D6 function or reduced gene expression, as outlined above. These patients are at increased risk of tamoxifen treatment resistance and, therefore, CYP2D6 variant testing would be recommended before treatment initiation to avoid exacerbation of the disease.

3.2 Induction Interactions

In general, gain-of-function mutations are less common than loss-of-function mutations. These rare kinds of mutations could result in increasing the expression of the gene via activating its transcription. This could result in increased gene products (RNA/proteins) leading to increased enzyme metabolic activity. Induction interactions occur when increased activity of the enzyme could reduce the efficacy of active drugs (as a result of increased deactivation of the drug) or increase the toxicity of prodrugs (as a result of increased production of active metabolites). Below, I present two examples explaining these two different scenarios.

The proton pump inhibitors (PPIs) lansoprazole, omeprazole, and pantoprazole are mainly metabolized by CYP2C19. Due to the convincing evidence from the literature that the increased activity variant *CYP2C19*17* is associated with decreased PPI plasma concentration and, in turn, increased chance for therapeutic failure, it is recommended by the CPIC to increase the daily dose of these agents for carriers of this variant allele [18].

The prodrug codeine, which is activated into morphine by CYP2D6 enzyme, has also received genotype-guided CPIC recommendations. They recommend avoiding codeine therapy and using an alternative analgesic agent for those with rapid and ultra-rapid CYP2D6 metabolism status due to the strong evidence of increased codeine toxicity [19].

4. Drug-Drug-Metabolizing Enzyme Gene Interactions (DDMEGIs)

4.1 Inhibitory Interactions

Inhibitory effects of drugs and genotype can alter substrate metabolism by both drug and genotype impacting on the same metabolizing enzyme, or on two distinct routes of metabolism.

In general, poor metabolizers are expected to experience the highest substrate drug plasma concentration, compared to other genotypes, when co-treated with inhibitors. For example, co-administration of simvastatin (a CYP2C9 inhibitor) with warfarin (a CYP2C9 substrate) has been shown to reduce warfarin dosage requirements in *CYP2C9*3* carriers by a greater percentage as compared to non-carriers (29% vs 5% respectively) [20].

A similar conclusion has been reported with celecoxib (Table 1, [21]). The inhibitory effect of drug and genotype is not always additive – genetically poor metabolizers may have only limited further enzyme inhibition by the administration of an inhibitory drug. For instance, a statistically significant elevation in rabeprazole (a CYP2C19 substrate) plasma levels was observed in both normal metabolizers and heterozygous genotype carriers after treatment with fluvoxamine (a *C*YP2C19 inhibitor) while no additional clinically significant elevation was detected with poor metabolizers who have already experienced the highest rabeprazole plasma levels [22]. A similar scenario is seen in other examples (Table 1, [23-26]).

Where a drug is metabolized by two or more CYP enzymes, then inhibition of one of these enzymes alone (by drug or genotype) may have minimal effect, due to redundancy of the pathways. However, if a genotype and interacting drug affect these different routes of metabolism, then the interaction may be very large. For example,

it has been observed that for voriconazole (a CYP2C19 & CYP3A4 substrate) bioavailability is increased markedly (~ 5.6-fold) in patients who have reduced CYP2C19 activity and are administered with atazanavir or ritonavir (potent CYP3A4 inhibitors) [27]. A similar scenario can be noted with other examples (Table 1, [28-30]).

Prodrugs, on the other hand, require the function of certain CYPs to be therapeutically active, and in these cases, the effect is the opposite to that described above. Clopidogrel, for example, is activated by CYP1A2, CYP2B6, CYP3A4, CYP2C9, and CYP2C19 [31]. Carriers of LOF variants in one or more of these genes and co-administered with their inhibitors are at increased risk for treatment resistance. For instance, carriers of *CYP2C19*2* and/or **3* alleles who are treated with clopidogrel and proton pump inhibitors (CYP2C19 inhibitors) were observed to be more likely to have reduced clopidogrel efficacy; the addition of a third risk factor (e.g., calcium channel blockers (CYP3A4 inhibitors)) was also correlated with a greater reduction in the efficacy of clopidogrel [32,33].

4.2 Induction Interactions

Increased-metabolism of active drugs by an enzyme inducer or gain-of-function variant will result in reduced efficacy of the victim drug. For example, when voriconazole (a CYP2C19 substrate) is co-prescribed with carbamazepine (a CYP2C19 inducer), the voriconazole dose is usually increased to overcome this increased metabolism. In a case report, therapeutic concentrations of voriconazole were not achieved, as the patient carried two gain-of-function *CYP2C19 *17* variants [34].

The opposite effect is seen with prodrugs. Increased metabolism by an enzyme-inducing drug or gain-of-function variant, will result in high plasma levels of active metabolites leading to increased side effects and/or efficacy. Thus, patients carrying *CYP2C19*17* gain-of-function variants have increased conversion of clopidogrel to active metabolites resulting in reduced cardiovascular events and/or increased bleeding episodes [35-44]. Co-administration of an inducer of CYP1A2, CYP2C9, and/or CYP3A4 would be expected to result in greater efficacy of clopidogrel, with increased risk of bleeding. However, no studies have been published to establish this.

4.3 Phenoconversion Interactions

As described above, a temporary phenotype shift can be seen when the perpetrator drug and genetic effect are opposed. For example, the presence of reduced function *CYP2C9* variants results in reduced tolbutamide (a CYP2C9 substrate) metabolism, yet co-treatment with rifampicin (a CYP2C9 inducer) in these patients reverses this genetic effect resulting in a two-fold increase in tolbutamide clearance [45].

Conversely, proton pump inhibitors (CYP2C19 inhibitors) treatment with clopidogrel results in phenoconversion in genetically determined ultra-rapid phenotype to a poor metabolizer status indicated by loss of clopidogrel efficacy [46].

The beneficial side of phenoconversion interactions is that genetically determined phenotypes can be normalized by the addition of medications of opposite effects on metabolism. For example, resistance to nortriptyline (a CYP2D6 substrate) due to abnormally rapid metabolism has been successfully reversed and normalized with the addition of paroxetine (a CYP2D6 inhibitor), which produces a recovery of nortriptyline therapeutic plasma levels [47].

The three figures below show the predicted changes of plasma levels of active drugs and active metabolites of prodrugs with and without the presence of inhibitors and/or LOF variants (Figure 1) and with and without the presence of inducers and/or GOF variants (Figure 2) and presents different scenarios of phenoconversion interactions (Figure 3).

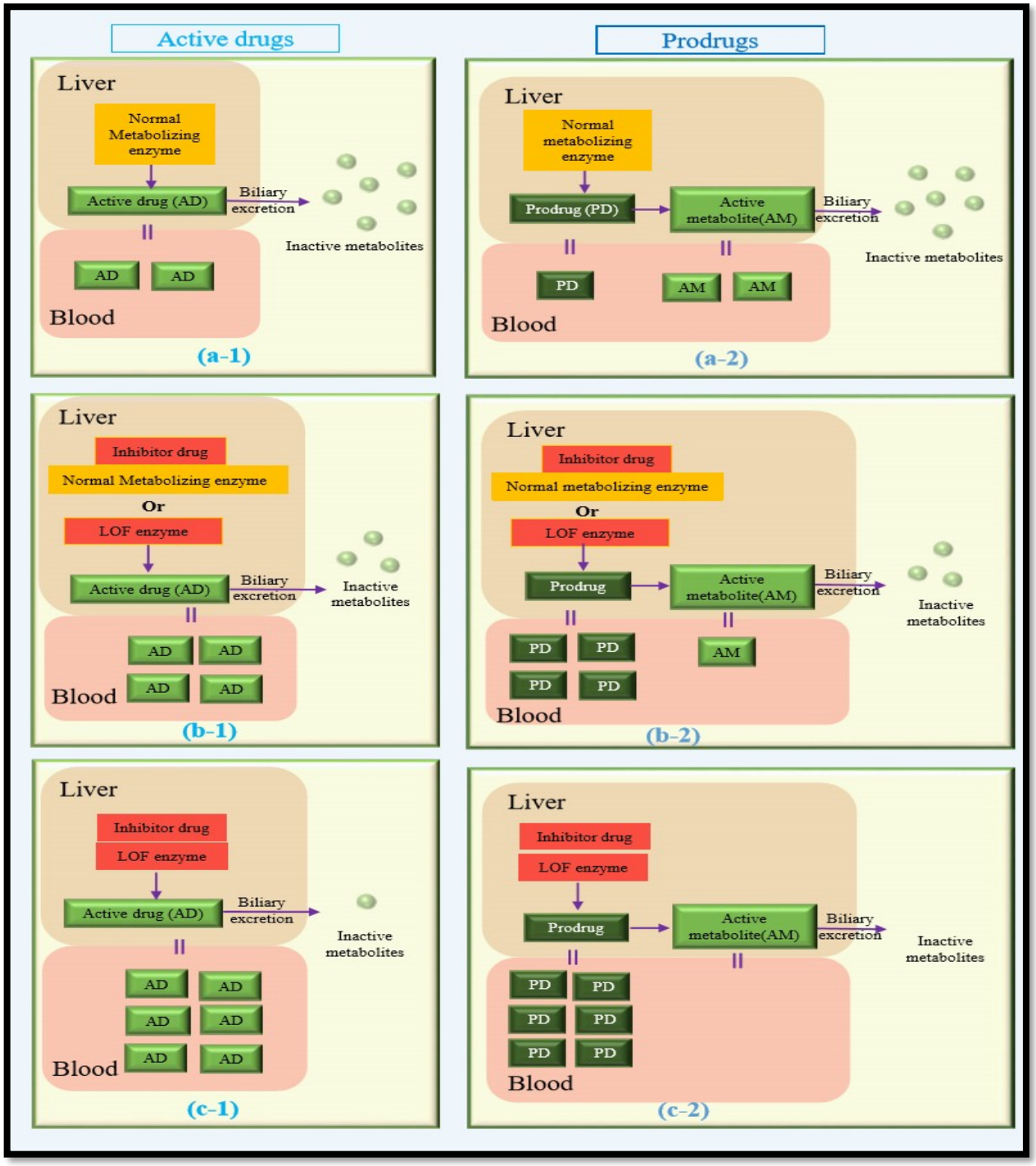


Figure 1: The predicted active drug /active metabolites of prodrugs plasma levels and biliary excretion changes without or with the presence of **inhibitors** or **LOF** variants or both on metabolizing enzymes. The predicted **active** drug / active metabolites of **prodrugs** plasma levels and biliary excretion changes without (a-1/a-2) or with the presence of inhibitors or LOF variants (b-1/b-2) or both (c-1/c-2) on metabolizing enzymes. (a-1/a-2) represent the normal scenario with no interacting drug or genetic variant. In b-1/b-2) either an inhibitory drug or loss of function variant (LOF) in the metabolizing enzyme, results in reduced metabolism to inactive metabolites , and increased (b-1)/decreased(b-2) active drug in the systemic circulation. In c-1/c-2) the presence of inhibitory drug and the LOF genetic variant combine to produce greater increase(c-1)/decrease(c-2) in the systemic concentration of active drug.

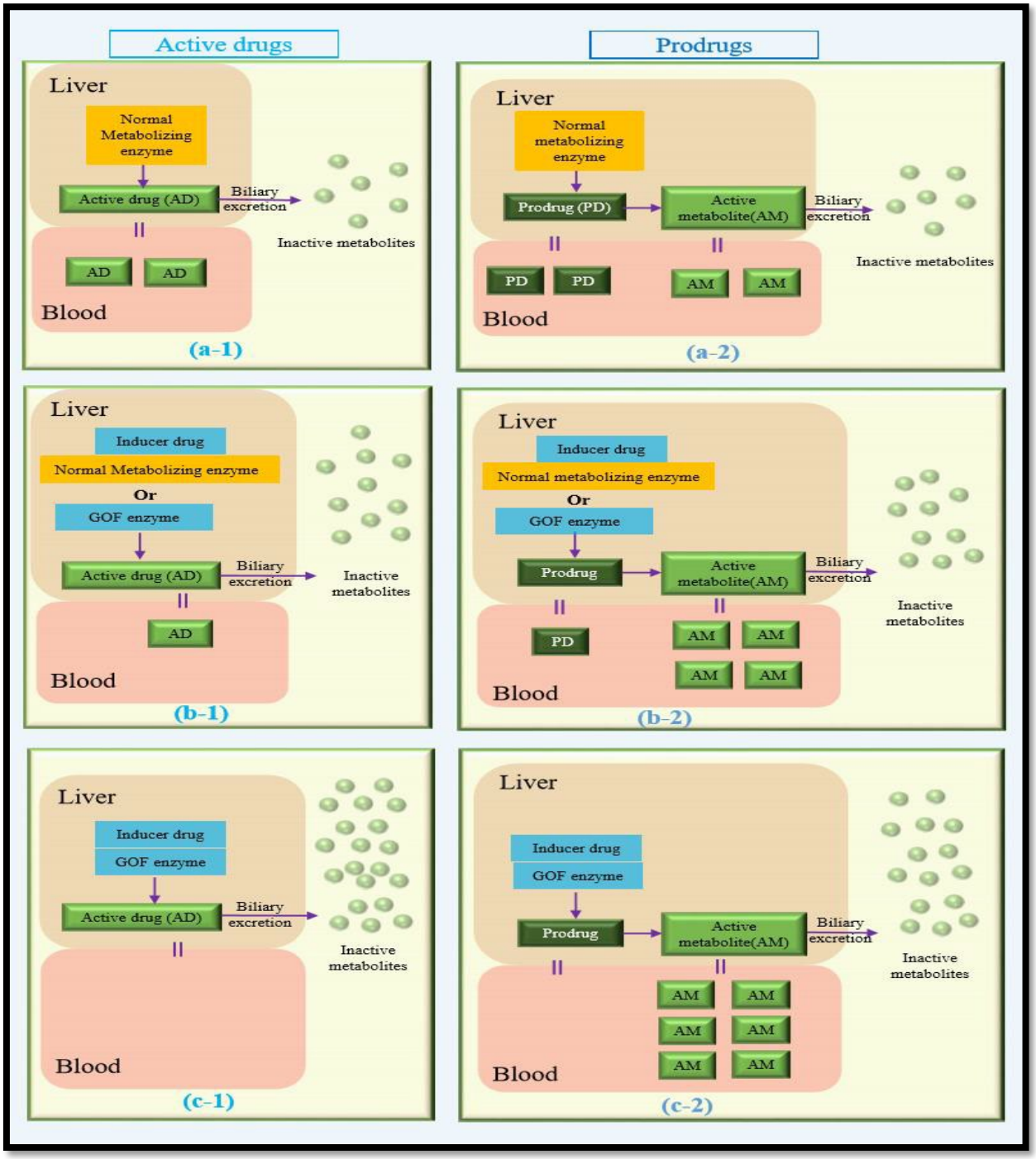


Figure 2: The predicted active drug / active metabolites of prodrugs plasma levels and biliary excretion changes without or with the presence of **inducers** or **GOF** variants or both on metabolizing enzymes. The predicted **active** drug / active metabolites of **prodrugs** plasma levels and biliary excretion changes without (a-1/a-2) or with the presence of inducers or GOF variants (b-1/b-2) or both (c-1/c-2) on metabolizing enzymes. (a-1/a-2) represent the normal scenario with no interacting drug or genetic variant. In b-1/b-2) either an inducer drug or gain of function variant (GOF) in the metabolizing enzyme, results in increased metabolism to inactive metabolites , and decreased(b-1)/increased(b-2) active drug in the systemic circulation. In c-1/c-2) the presence of inducer drug and the GOF genetic variant combine to produce greater decrease(c-1)/increase(c-2) in the systemic concentration of active drug.

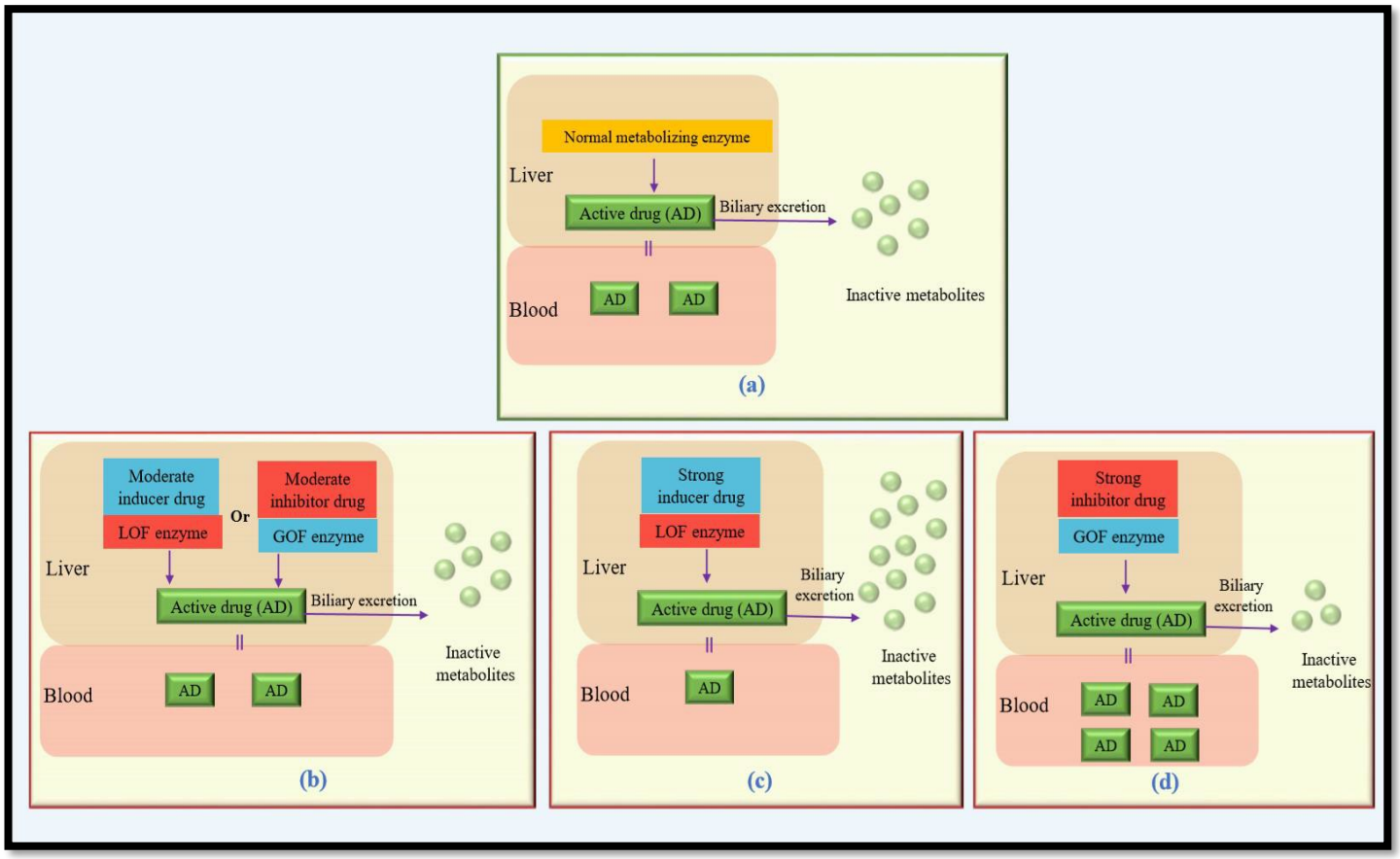


Figure 3 : Different scenarios of **phenoconversion** interactions where genetic effects may be reversed or shifted in the opposite direction . (a) represents the normal scenario with no interacting drug or genetic variant. In b) the effect of loss of function variant (LOF) or gain of function variant (GOF) is reversed with the presence of a moderate inducer drug or a moderate inhibitor drug respectively and results in a clinical outcome similar to the normal situation (a). In c) the presence of a strong inducer drug has temporarily shifted a poor metabolism status into a rapid metabolism status and results in decreased active drug in the systemic circulation. In d) the presence of a strong inhibitor drug has temporarily shifted a rapid metabolism status into a poor metabolism status and results in increased active drug in the systemic circulation.

5. Transporter-related interactions

Drug transporters govern the movement of pharmaceutical compounds from and into different body tissues. The liver, kidney, blood-brain barrier (BBB), and intestine are the key sites of transporters that influence drug pharmacokinetics. In addition to summarizing the distribution and localization of transporters, Figure 4 below (which has been formulated with the aid of reference [48]) also classifies transporters into three categories according to the similarity of transport directions in different tissue types: Group I efflux transporters, Group II efflux transporters, and Group III (uptake) transporters. The following paragraphs discuss these three categories of transporters and their drug-related interactions.

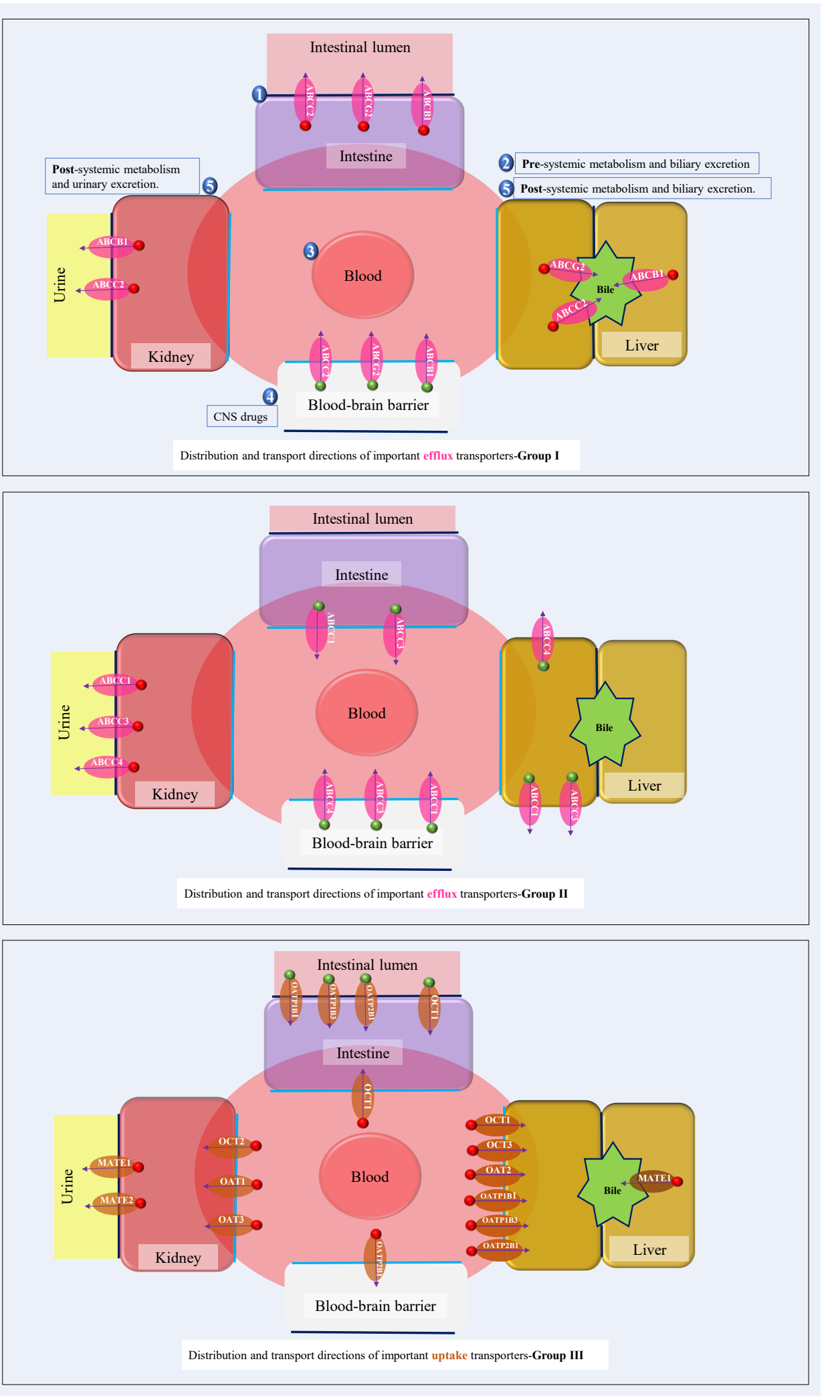


Figure 4: Drug transporters as classified into three categories according to the similarity of the transport directions in different tissue types.

Numbers from 1 to 5 = order of oral drug movement through different tissue types. Non-oral drug formulations bypass the effect of intestinal transporters.

/ = **increased** / **decreased** substrate drug plasma level is predicted as a result of impairment of this transporter due to LOF variants or inhibitors. The reverse is predicted with GOF variants or inducers.

The presence of the two factors (i.e. LOF variant+ an inhibitor or GOF variant + an inducer) is predicted to double the clinical impact with neutralizing or shifting the clinical effect when the preparator drug and genetic effect are opposed (photoconversion interactions).

= Apical membrane
= Basolateral membrane

Information about distribution and transport directions has been quoted from reference number 48

This page was left intentionally blank

5.1 Group I efflux transporters

P-glycoprotein 1 (P-gp, ABCB1), multidrug resistance-associated protein 2 (MRP2, ABCC2), and breast cancer resistant protein (BCRP, ABCG2) transporters are expressed in the intestine, liver, kidney, and blood-brain barrier (BBB), sharing similar transport pathways. They efflux substrates back to the intestinal lumen, facilitate hepatic and renal excretion (excluding BCRP), and work inversely in the BBB where they protect the brain from the entry of xenobiotics and return them back to the systemic circulation. Blocking their function in the intestine, liver, or kidney is expected to elevate a substrate's systemic exposure (although opposite effects would be predicted if inhibiting transport across the BBB). In the sections below, I show examples of DGIs and DDGIs related to group I transporters.

5.1.1 Drug-Transporter Gene Interactions (DTGIs)

The rs1045642 (C>T) variant is the most common studied SNP in ABCB1 transporter. Pharmacogenomic studies report opposing effects of this SNP with different drugs. For example, carriers of this variant experienced low methotrexate plasma levels and toxicity [49], while this variant shows increased hepatotoxicity among nevirapine users [50]. Both of these two findings have been classified under the same category of high evidence findings (level 2A) according to the Pharmacogenomics Knowledge Base (PharmGKB) which has also reported many other similar examples with other drugs [51].

According to PharmGKB, one of the best evidence available to date for ABCG2 transporter is the correlation between rs2231142 (G>T) variant and rosuvastatin efficacy. Carriage of the T allele has been associated with increased rosuvastatin lipid-lowering efficacy [52]. This is believed to occur as a result of increased

rosuvastatin accumulation in the liver, the site of rosuvastatin action, due to decreased hepatic ABCG2 efflux activity.

To date, there are no published genotype-based clinical guidelines for any drug regarding genetic variants under this group of transporters.

5.1.2 Drug-Drug-Transporter Gene Interactions (DDTGIs)

In this group, the most evidence for DDTGI comes from drugs altering ABCB1 (P-gp) transport and genetic variants in the gene encoding this transporter. For example, cyclosporine is an ABCB1 substrate. Diltiazem (a moderate ABCB1 inhibitor [53]) has been shown to increase cyclosporin trough concentrations in Chinese patients who carry the TT genotype (low P-gp activity) at rs1045642(C>T) in ABCB1; yet no effect was seen in other *ABCB1* genotypes (e.g., CC at rs1045642) [54]. Methadone is also a P-gp substrate, acting in the brain and effluxed across the BBB via P-gp. Patients with the TT genotype at rs1045642 and treated with quetiapine (an ABCB1 inhibitor) experienced the lowest increase in methadone plasma levels compared to those with CT or CC genotypes (3% vs 23% vs 33% respectively) [55]. Low methadone plasma levels in this study would be explained by loss of the ABCB1 protective function in the BBB, which results in an increased intracerebral concentration of this central nervous system (CNS) drug. As a result of a similar DDTGI mechanism, the CNS drug granisetron was associated with increased efficacy in Japanese subjects (Table 1, [56])

In some cases, it seems that adding strong inhibitors abolishes the effect of genotype. For example, no additional inhibitory effects were detected in carriers of different genotypes of the rs1045642 (C>T) *ABCB1* variant who were either on

dabigatran/rivaroxaban-clarithromycin combination or tacrolimus-itraconazole combination (ABCB1 substrates-ABCB1 strong inhibitors [53])[57,58].

5.2 Group II efflux transporters

MRP1 (ABCC1), MRP3 (ABCC3), and MRP4 (ABCC4) share the similar transport direction in the kidney and BBB as the Group 1 transporters. However, in the liver, they are expressed in the basolateral membrane working to pump drugs back into the systemic circulation. Unlike group 1 transporters, there are no published studies describing DDTGIs for group II transporters. So here I report DGTIs to highlight the potential mechanisms whereby genes and drugs that alter these transporters may influence drug outcomes.

MRP1, for example, transports the active metabolite of irinotecan (SN-38) out of hepatocytes into the blood contributing to the well-known side effect of irinotecan induced neutropenia [59]. The reduced function variant, rs17501331, in the *ABCC1* gene is associated with a low incidence of neutropenia; the reverse effect was detected with the gain-of-function variant rs6498588 in the same gene [60]. In some cases, reduced ABCC1 activity could be disadvantageous. For instance, carriers of the AA genotype (suggested to be linked with reduced ABCC1 activity) at rs246240 (A>G) *ABCC1* variant are at higher risk for developing methotrexate gastrointestinal and hepatic toxicity as a result of increased methotrexate intracellular accumulation [61]. Of note, MRP1 is also expressed in the myocardium protecting the heart from the entry of xenobiotics [62]. For example, the reduced transport associated with the rs45511401 (G>T) variant in *ABCC1* gene increases the chance of developing cardiotoxicity due to intracellular accumulation of doxorubicin [63].

ABCC1 and ABCC3, in contrast to ABCB1, ABCC2 and ABCG2, are expressed in the basolateral membrane of the intestine effluxing substrates into the portal

circulation. As orally administered drugs are first exposed to intestinal transporters, any modification of their role might affect drug concentration in the other tissues (liver, kidney, or BBB). C.1037C>T and c.1820G>A *ABCC3* variants, for example, have low transport activity [64] suggesting their potential to diminish the bioavailability of oral ABCC3 substrates irrespective of subsequent alteration in transport into other tissues, or subsequent metabolism.

5.3 Group III (uptake) transporters

In the liver, kidney, and BBB, all important uptake transporters (organic cation transporters (OCTs)1/2/3, organic anion-transporting polypeptide (OATP) 1B1/1B3/2B1, and multidrug and toxic compound extrusion proteins (MATE) 1/2), follow an identical primary route for transporting their substrates: from the systemic circulation into different tissues or urine/bile in case of MATEs. Consequently, reducing or increasing these transport capacities would result in increased or reduced systemic drug concentrations, respectively. The reverse effects are seen with the uptake transporters expressed in the intestinal apical membrane such as OATPs and OCT1 since the transportation pathway is in the opposite direction. In the sections below, I show examples of DGIs and DDGIs related to group III transporters.

5.3.1 Drug-Transporter Gene Interactions (DTGIs)

In many situations, the efficacy of a drug relies upon the ability of that drug to access certain tissues. Statins are taken up into the liver by OATP1B1(SLCO1B1), and this is crucial for their lipid-lowering effect. Reducing this uptake pathway reduces statin efficacy and raises plasma concentrations, resulting in myopathy and, rarely, rhabdomyolysis. The rs4149056 (T>C) (*SLCO1B1*15*) variant has been widely studied, and in 23 studies [65-87], this variant has been persistently connected to

increased statin plasma exposure, muscle aches, dose reduction, and/or treatment-resistant phenotypes. The strongest evidence to date is available between simvastatin and rs4149056 (T>C) *SLCO1B1* variant in term of increased myopathy risk. Therefore, CPIC guidelines recommend to reduce the dose or use an alternative statin not significantly affected by SLCO1B1 for carriers of this variant allele [88]. Of note, this is the only genotype-phenotype association which has received CPIC recommendations at the level of drug transporters available to date.

5.3.2 Drug-Drug-Transporter Gene Interactions (DDTGIs)

In some circumstances, altering the uptake transporter function can increase ADRs. For example, it has been observed that carriers of two *OCT1* (*SLC22A1*) reduced function alleles who were treated with OCT1-inhibitors were over four times more likely to develop gastrointestinal side effects with metformin (an OCT1 substrate) treatment, which would be attributable to metformin accumulation in the intestinal lumen (assuming apical OCT1 localization) [89]. This finding was supported by a previous study [90]. At the level of renal uptake transporters, other DDTGIs have been reported in which carrying the mutant alleles, and the co-administration of inhibitors was linked to increased metformin plasma levels/toxicity or reduced clearance (see Table 1, [91,92]). By contrast, reducing transport in some cases may reduce certain side effects. For instance, cisplatin (an OCT2 (SLC22A2) substrate) is both a nephrotoxic and an ototoxic agent. People carrying the rs316019 (C>A) *OCT2* mutation were protected from these adverse reactions as the variant resulted in a reduced transport of cisplatin into the kidney and the inner ear (cochlea) (where OCT2 is expressed as well) [93-95].

A number of other DDGIs have been described for the SLCO1B1 transporter. For example, although the increase in pravastatin (a SLCO1B1 substrate) AUC after treatment with ritonavir (a SLCO1B1 inhibitor) was not statistically significant (21% increase vs pravastatin alone), a large interaction was seen in those carrying the *SLCO1B1*15* or **17* haplotypes, with a resulting 113% elevation in pravastatin AUC [96]. Other DDTGIs with a similar mechanism have also been published (see Table 1, [97-99]). Interestingly, unlike the ritonavir example just outlined, in some situations reduced function variants do not show any significant PK change until after the addition of inhibitors. For example, patients with AG or AA genotypes at rs2289669 (G>A) of the MATE1 transporter only had significantly lower metformin (a MATE1 substrate) clearance compared to carriers of GG genotype after treatment with ranitidine (a MATE1 inhibitor) [100].

6. DGIs/DDGIs and challenges in clinical practice

Metabolizing enzyme and transporter substrates, inducers, or inhibitors are not fully documented in many popular drug interaction databases, leaving physicians unaware of potentially important interactions. In addition, most of the resources commonly used by prescribers (e.g., Stockley's, Micromedex, Drug.com, RxList, or other drug interaction checkers) do not consider genetic variation when classifying drug interactions into minor, moderate, or major classes. Genetic variation may markedly increase or ameliorate the severity of potential drug interactions and do need to be considered when considering real-world use of drugs.

In this introduction, I have discussed the different mechanisms of interactions in their simplest forms with the assumption that the patient is free of transporter polymorphisms or inhibitors/inducers in the case of discussing DMEGIs/DDMEGIs and vice versa with DTGIs/DDTGIs. However, in real-world clinical practice, achieving precisely tailored drug therapy requires a detailed examination of all mutations in the candidate enzyme or transporter genes with a good awareness of all prescribed medications and possible pathways of interaction. Thus, the clinical scenario ranges from a relatively simple picture where the effect of genotype and interacting drug(s) can be approximated and treatment altered accordingly, to a far more complex scenario where physiologically-based PK (PBPK) modelling may be helpful and where the evaluation of large-scale clinical data linked to genotypes is required to evaluate the clinical impact of multiple interacting drugs/multiple genotypes on drug outcomes.

Consider a relatively simple scenario: a patient with type 2 diabetes treated with metformin (has no effect on CYPs) who carries reduced function variants in *CYP2C9* (**2* or **3* variants) and who is started on gliclazide (a CYP2C9/19 substrate). Reduced metabolism of gliclazide will result in increased efficacy [101] and increased risk of hypoglycemia [102]. The metformin use will not alter this DGI. However, if this patient were also treated with pioglitazone and/or atorvastatin (both are CYP2C9/19 inhibitors), they would be at potentially even greater risk of gliclazide-induced hypoglycemia and should be treated with a reduced dose of gliclazide. However, even for this simple scenario, such DDGI studies have not been reported; nor have dosing algorithms been developed to date for patients with *CYP2C9* variants prescribed sulphonylureas and as such it is difficult to implement this into drug interaction calculators.

There are many more complex scenarios where, for example, a combination of both metabolizing enzyme and transporter LOF/GOF variants, as well as inhibitors/inducers are included. This kind of interaction may be only initially predictable when all their sub-interactions result in the same clinical effect. For instance, reduced CYP3A4 and SLCO1B1 activities can both result in increased AUC of the substrate drug, and greater harm would be anticipated. Carriers of the TC genotype of *SLCO1B1* rs4149056 (T>C) variant who are treated with amlodipine (a CYP3A4 inhibitor) experienced a 90% increased simvastatin AUC compared to subjects not treated with amlodipine and wild-type for rs4149056 [103]. A similar scenario was reported with other two case reports (see Table 1, [104,105]).

In other situations, sub-interactions do not share a similar clinical effect. Here, predicting the overall clinical outcome is challenging. As an illustration, oral rosuvastatin is mainly eliminated via biliary excretion with a minor contribution of

CYP2C9 to its metabolism [106]. This implies that its transporters (e.g., ABCC2, ABCG2, ABCC1, and SLCO1B1) are the core players in its elimination. The concomitant administration of verapamil (an ABCC1/2 inhibitor) and venlafaxine (an ABCG2 inducer) in those who have inherited *CYP2C9**3 and/or *SLCO1B1* rs4149056 (T>C) LOF variants results in unpredictable clinical consequences. CYP2C9, SLCO1B1, and ABCC2 impairment would boost rosuvastatin AUC, inducing ABCG2 would lower rosuvastatin AUC, and inhibition of ABCC1 could result in both increase or decrease in AUC (high AUC if the site of interaction is in the kidney and low AUC if it is in the intestine or liver).

The exact estimation of the predicted net AUC following a certain drug interaction relies on calculating the contribution of each metabolizing enzyme and transporter to the elimination process (i.e., the degree of sensitivity of substrates), inhibition/induction potency of the perpetrator agent or the net effect of multiple inhibitors, inducers, or both, and the net percentage of reduction/elevation in the enzyme/s and/or transporter/s activity caused by a single or more SNPs. The outcome of such a hugely complex scenario is impossible to predict by the clinician and requires a clinical support tool based upon a PK DDGI prediction algorithm. Most of the current work concentrates on generating DD interaction predictors rather than the combined effect of both drugs and variants. However, using PBPK models, one predictor tool (https://www.ddi-predictor.org/) has recently been successfully generated to estimate drug exposure and the recommended dose following the dual action of both the perpetrator drug and mutations in certain CYPs (CYP2D6, CYP2C9, and CYP2C19) [107]. Other PBPK models do attempt to incorporate genotype and drug-drug interactions, but these do not model transporter variants well and have yet to translate through into clinically useful tools [108].

An alternative method to evaluate the impact of DGIs/DDGIs is via measurement of endogenous biomarkers of metabolizing enzymes and transporters function rather than plasma concentrations of substrate drugs. Multiple enzymes/transporters-related biomarkers have been identified and are summarized in a published review [109] from which I highlight some examples in this paragraph. For instance, it has been shown that the cholesterol, cortisone, and cortisol metabolites: 4β-hydroxycholesterol, 6β-hydroxycortisone, and 6β-hydroxycortisol respectively, which are catalyzed by CYP3A4 activity, are increased under the effect of inducers and decreased with inhibitors of CYP3A4. It was also recognized that bufotenine is a major metabolite resulting from the metabolizing activity of CYP2D6. With regard to transporters, several studies have observed the association between increased bilirubin plasma levels and reduced hepatic OATP1B1/1B3 uptake function. The similar scenario was noted recently with the novel biomarkers coproporphyrins I and III (CPs I and III) where plasma CPs levels elevated with the inhibition of these transporters to a similar extent as with rosuvastatin. In DGIs/DDGIs studies, endogenous biomarkers can be utilized to predict the effect of genetic variants and/or inhibitors/inducers on the substrate drugs plasma levels.

It is worth noting that potential DDIs do not necessarily reflect actual interactions. It has been observed that clinically significant interactions are consistently lower than theoretically predictable interactions [110]. However, the authors noted that 20% of ADRs are linked with DDIs; most of them are serious, with a high percentage of fatal cases. They also concluded that therapeutic failure secondary to DDIs, which is usually underestimated, represents a considerable part of total DDIs-related undesirable effects. The degree of clinical significance can be judged by observing other risk factors associated with a potential DDI such as polypharmacy and genetic

variants. Polypharmacy is commonly seen with elderly and hospitalized patients making them the most vulnerable patient's sub-groups to clinically significant interactions besides carriers of risky genetic variants. In addition, not all types of DGIs/DDGIs are expected to be common. Induction and phenoconversion interactions are predicted to be seen with lower incidence compared to inhibitory interactions as the majority of perpetrator drugs are inhibitors rather than inducers and most of the functional genetic variants are loss rather than gain-of-function mutations.

The increasing availability of 'big data' linking health data and genomics has the potential to evaluate the real-world clinical impact of multiple drugs/multiple variant interactions. A number of data sets are now available or about to become available for study. In Scotland national prescribing and linked outcomes are available for the entire population enabling evaluation of real-world DDIs, and with an increasing bioresource (https://www.registerforshare.org) it should be possible to evaluate DGIs/DDGIs in ~500K people over the next few years. In addition, other resources such as UK biobank including genetic information on 500K individuals (with primary care data available on 200K during 2019) and other national bioresources (such as the Danish biorepository) and US bioresources linked to EHRs (EMERGE network) will enable the evaluation of n-way DDG interactions to identify clinically important interactions that can be incorporated into clinical decision support tools in the future.

7. Conclusion and the aim of this PhD project

Dozens of new pharmaceutical compounds enter the market each year, and a considerable number of patients are prescribed multiple drugs that necessitate the utilization of drug interaction databases for better management. One of the major limitations of these drug interaction checkers is the omission of the genetic effect on drug interactions. This reflects both the lack of clinical studies that quantify potential DGIs/DDGIs and the fact that genetic information is rarely available on patients at the point of prescribing. In this introduction, I have illustrated, with some examples, various mechanisms by which DGIs/DDGIs can occur at the level of metabolizing enzymes, drug transporters, or both. As DDGIs are a relatively new topic in pharmacogenomics with very limited publications, I have established a detailed classification framework of these interactions based on clinical studies or case reports. I have also shown the different degrees of complexity clinicians may face in judging the predicted clinical outcome following a certain DDGI. The more factors that are included, the more challenging it becomes to evaluate the outcome. There is a need for PBPK models, clinical studies and real-world evaluation of drug outcomes linked to genetic information to develop clinical useful DDGI models, to reduce adverse DDIs and improve drug outcomes in the setting of increasing multi-morbidity and polypharmacy.

Due to the lack of DG and DDG interactions studies, my aim in this PhD project is to investigate a large variety of drugs and drug combinations and study the impact of different genetic variants in genes known to affect drug pharmacokinetics on different drug response phenotypes with the aim of uncovering novel DG/DDG associations. I will begin by describing the general methodology in the next chapter before

presenting our findings from DG and DDG interaction studies in chapters III, IV, and V.

Table 1: classification of drug-drug-gene interactions according to their different mechanisms as observed from findings of clinical studies or case reports.

Metabolizing enzymes interactions			
Inhibitory interactions-the victim drug as an active drug			
Interaction type	Interaction model	Example/s	Clinical outcome
Single enzyme interactions (Both the inhibitor and the genotype affect the same enzyme)	Substrate + inhibitor +LOF variants in the same enzyme inhibited by the inhibitor	Warfarin+simvastatin+*CYP2C9*3* in *CYP2C9* gene	*CYP2C9**3 carriers require lower warfarin dose compared to non-carriers [20].
		Lansoprazole/rabeprazole+ fluvoxamine+ poor metabolism genotype in *CYP2C19* gene.	Although the greater increase in the substrate drug plasma level is always seen with poor metabolizers, the effect of adding inhibitors is observed more clearly in normal metabolizers followed by heterozygous genotype carriers with the lowest effect being with poor metabolizers as they already carry reduced activity metabolizing enzymes [22-26].
		Omeprazole+ moclobemide+ poor metabolism genotype in *CYP2C19* gene.	
		Metoprolol+ diphenhydramine+ poor metabolism genotype in *CYP2D6* gene.	
		Metoprolol+ dronedarone+ poor metabolism genotype in *CYP2D6* gene.	
	Substrate +2nd substrate (competitive inhibitor) +LOF variants in the same enzyme inhibited by the inhibitor	Warfarin + celecoxib+ *CYP2C9*2/*3* in *CYP2C9* gene	Increased bleeding risk with warfarin in those carrying the variant *CYP2C9* alleles [21].
Multiple enzymes interactions (The inhibitor and the genotype affect different enzymes)	Substrate +inhibitor +LOF variants in an enzyme differ from the one inhibited by the inhibitor	Voriconazole + atazanavir/ritonavir (CYP3A4 inhibitors) + poor metabolism genotype in *CYP2C19* gene.	Poor metabolizers experienced marked AUC increase from the substrate drugs compared to normal metabolizers [27-30].
		Tacrolimus + voriconazole (a CYP3A4 inhibitor) + poor/intermediate metabolism genotype in *CYP2C19* gene.	
		Voriconazole + erythromycin (a CYP3A4 inhibitor) + poor metabolism genotype in *CYP2C19* gene.	
		Lansoprazole + clarithromycin (a CYP3A4 inhibitor) + poor metabolism genotype in *CYP2C19* gene.	

Inhibitory interactions-the victim drug as a prodrug			
Single enzyme interactions	Substrate +inhibitor +LOF variants in the same enzyme inhibited by the inhibitor	Clopidogrel + proton pump inhibitors (CYP2C19 inhibitors) + *CYP2C19*2/*3* in *CYP2C19* gene	Poor metabolizers were more likely to develop clopidogrel resistance compared to normal metabolizers [32].
Multiple enzymes interactions	Substrate + 2 inhibitors of 2 different enzymes +LOF variants in one of the enzymes inhibited by one of the two inhibitors.	Clopidogrel + proton pump inhibitors (CYP2C19 inhibitors) + calcium channel blockers (CYP3A4 inhibitors) + *CYP2C19*2/*3* in *CYP2C19* gene	Poor metabolizers developed a marked reduction in clopidogrel efficacy compared to normal metabolizers [33].
Induction Interactions			
Single enzyme interactions	Substrate +inducer +GOF variant in the same enzyme induced by the inducer	Voriconazole + carbamazepine+ *CYP2C19*17/17* in *CYP2C19* gene	Major loss of voriconazole efficacy was seen in one patient [34].
Phenoconversion Interactions			
Single enzyme interactions	Substrate +inhibitor+ GOF variant in the same enzyme inhibited by the inhibitor	a) Ultra-rapid metabolism > poor metabolism Clopidogrel+ proton pump inhibitors+ *CYP2C19*17* in *CYP2C19* gene	Loss of clopidogrel efficacy [46].
		b) Ultra-rapid metabolism > normal metabolism Nortriptyline + paroxetine + ultra-rapid metabolism genotypes in *CYP2D6* gene.	Normal nortriptyline therapeutic plasma levels [47].
	Substrate +inducer+ LOF variant in the same enzyme induced by the inducer	c) Poor metabolism> ultra-rapid metabolism Tolbutamide + rifampin+ intermediate or poor metabolism genotypes in *CYP2C9* gene	Twofold increase in tolbutamide clearance [45].

Transporters interactions			
Inhibitory interactions-efflux transporters			
Single transporter interactions	Substrate/s +inhibitor/s +LOF variant in the same transporter inhibited by the inhibitor	Cyclosporine + diltiazem + rs1045642(C>T) in *ABCB1* gene	TT genotype carriers at rs1045642 variant show increased cyclosporin trough concentration (no effect was detected with other genotypes) [54] and the lowest increase in methadone plasma level compared to other genotypes [55].
		Methadone + quetiapine + rs1045642(C>T) in *ABCB1* gene	
		[Granisetron + dexamethasone] + [doxorubicin+ cyclophosphamide] + rs2032582(G>T)* in *ABCB1* gene **G, but not T, allele was linked to reduced ABCB1 activity in this study with Japanese population.*	GG carriers show the highest response from the substrates compared to other genotypes [56].
		Dabigatran/rivaroxaban + clarithromycin **or** Tacrolimus + itraconazole + rs1045642 (C>T) in *ABCB1* gene	No effect was seen with any genotype [57,58].
Inhibitory interactions-uptake transporters			
Single transporter interactions	Substrate +inhibitor +LOF variant in the same transporter inhibited by the inhibitor	Metformin + OCT1 inhibitors (tricyclic antidepressants, citalopram, proton pump inhibitors, verapamil, diltiazem, doxazosin, spironolactone , clopidogrel , rosiglitazone , quinine , tramadol ,and codeine) + R61C, C88R , G401S, M420del, and G465R in *OCT1* gene.	Carriers of two LOF alleles were over four times more likely to develop metformin intolerance [89].
		Metformin + trimethoprim (a MATE1/OCT2 inhibitor) + rs2289669 in *MATE1* gene **or** rs316019 in *OCT2* gene	Increased metformin plasma concentrations [91].
		Metformin + cimetidine + c.808G>T in *OCT2* gene.	TT genotype carriers show the lowest rate of metformin clearance compared to other genotypes [92].

		Pravastatin + ritonavir + *SLCO1B1*15* or **17* haplotypes in *SLCO1B1* gene.	Carrying the variant alleles has elevated pravastatin AUC by 113%. No significant interaction was seen with normal genotypes. [96].
		Olmesartan + pravastatin **or** Repaglinide + gemfibrozil + rs4149056 (T>C) in *SLCO1B1* gene.	CC genotype carriers displayed higher plasma concentrations from the substrate drugs compared to other genotypes [97,98].
		Atorvastatin + rifampicin + rs4149056 (T>C) in *SLCO1B1* gene.	Although no significant difference in atorvastatin plasma concentration between the three genotypes was detected, a clear increase can be observed with CC genotype carriers and on the drug combination compared to those on atorvastatin only [99].
		Metformin + ranitidine + rs2289669(G>A) in *MATE1* gene	AG and AA genotype carriers experienced lower metformin clearance compared to GG genotype carriers [100].
Enzyme/s + transporter/s interactions			
Multiple enzyme/s and transporter/s interactions (The perpetrator drug and the genotype affect both enzyme/s and transporter/s of the same substrate)	Substrate + transporter/s and/or enzyme/s inhibitor/s and/ or inducers and/or 2nd substrate + LOF/GOF variants in one enzyme or more and/or in one transporter or more		
Simple (predictable) interactions (sub-interactions result in the same clinical outcome)			
		Simvastatin + amlodipine (a CYP3A4 inhibitor) + rs4149056 (T>C, LOF) in *SLCO1B1* gene.	Subjects with TC genotype experienced a 90% increase in simvastatin AUC compared to those not treated with amlodipine and carry normal *SLCO1B1* genotype [103].

	Fluvastatin + telmisartan (an ABCC2 inhibitor) + *CYP2C9*3* and -24C > T (LOF variants) in *CYP2C9* and *ABCC2* genes respectively.	An elevation in creatine kinase level was observed in a patient on this drug-drug-genotypes combination [104].
	Atorvastatin + pantoprazole (a SLCO1B1 and an ABCB1 inhibitor, and a CYP3A4 substrate) + CC genotype* at rs4149056 (T>C) in *SLCO1B1* gene + TT genotype* at rs1045642 (C>T) in *ABCB1* gene. ** CC and TT genotypes are associated with reduced activity.*	A patient on this drug-drug-genotypes combination has developed rhabdomyolysis and acute renal failure [105].
Complex (unpredictable) interactions (sub-interactions result in different clinical outcomes)		
	Rosuvastatin + verapamil (an ABCC1/2 inhibitor) and venlafaxine (an ABCG2 inducer) + *CYP2C9*3* and/or rs4149056 (T>C) LOF variants in *CYP2C9* and/or *SLCO1B1* genes respectively. *This example has been used for illustration purpose only (no studies were reported).*	Advanced PBPK prediction tools are required to be developed with conducting real-world clinical studies to address and understand this kind of complex interactions.

This page was left intentionally blank

Chapter II:

General Methodology

1. Introduction

This chapter will provide a general overview of databases and statistical methods utilized throughout this project as well as presenting the different methods used to define drug response/intolerance phenotypes.

2. The clinical databases

We have used four databases to produce the main findings in this project which are each described in brief.

2.1 The UK biobank (UKBB) database

The information below has been summarized from references 111 and 112.

2.1.1 Participants and data collection

The UKBB is a dataset of ~ half a million participants. They were recruited in the period between 2006 and 2010 when their age range was between 40 and 69 years. Participants were assessed in 22 centres across the UK, including a mix of different ethnicities. A large variety of phenotypic data was collected during these assessment visits. Sociodemographic, environmental, lifestyle, physical, medical history, and treatments information were collected via questionnaire and interviews. At the time of starting this PhD in March 2017, the available data were cross-sectional until September 2019 when longitudinal primary care data became available for ~ 230,000 participants. All participants provided consent to use their anonymized medical data for relevant research. Researchers are not required to submit a separate ethics approval to use the data as the UKBB has its own independent Ethics and Governance Framework which formulates its policies and access arrangements.

2.1.2 Genotype data

Genotype data for ~150,000 participants were first released in 2015 before the availability of genotype data for most of the cohort (n~ 488,377) in 2017 at the time of starting this project. Two novel closely-related genotyping arrays were specifically developed for the UKBB genotyping project: Applied Biosystems™ UK BiLEVE Axiom™ Array by Affymetrix and Applied Biosystems™ UK Biobank Axiom™ Array. The former was used to genotype 49,950 individuals participating in the UK Biobank Lung Exome Variant Evaluation (UK BiLEVE) study, and the latter was used for genotyping the remaining 438,427 subjects with a total of ~825K markers included. Coding markers of all different MAF categories were included (i.e. MAF < 1% [rare], MAF =1-5% [low frequency], or MAF > 5% [common]). Most importantly, around 45K variants associated with different chronic diseases and pharmacogenomics were also included on the genotyping array.

Affymetrix used a GeneTitan® MultiChannel (MC) Instrument to analyze DNA samples, and a number of quality checks were performed. Samples that failed to meet these checks were excluded. Some markers were not previously assayed by Affymetrix and required the development of new assay methods. As these new attempts were not successful, this group of markers were excluded. This resulted in genotyping of 812,428 different variants.

A number of approaches were taken into account to examine the quality of genotyping by the UKB research team such as checking HWE deviation, sex effects, plate effects, array effects, batch effects, and the DNA sample quality. To reduce the effect of population stratification (as the UKBB population consists of 94.23% White, 1.94% Asian, 1.57% Black, 0.31 % Chinese, 0.58 % mixed, and 1.38% unknown

ethnicity), principal component analysis (PCA) was undertaken to control for population stratification effect. White British participants represent the majority of the UKBB cohort and can be used by researchers interested in a homogenous ethnicity to avoid the risk of ethnic diversity on HWE deviation. Markers are excluded when there is sufficient evidence of poor genotyping quality. The final number of variants becomes 805,426 for 488,377 participants.

For the purpose of increasing the number of tested markers, genotype imputation was carried out by the UKBB research team. Imputation is the process of predicting indirectly assayed genotypes for a sample of subjects. The imputation process includes pre-phasing the directly genotyped markers first before imputation. In the first step, markers which failed quality control for more than 1 batch, had > 5% missing data, or had a minor allele frequency of < 0.0001 were removed. Some samples which were recognized as outliers were also removed. This reduced the number of markers into 670,739 and reduced the number of samples into 487,442.

As the accuracy of imputation increases with the increased number of haplotypes in the reference panel and if there is a good match between the ancestries of the sample haplotypes and the reference panel haplotypes, the Haplotype Reference Consortium (HRC) panel was utilized. This panel consists of many more haplotypes with the European (including the British) ancestry representing the majority of individuals as the UKBB cohort. The HRC panel also consists the 1000 Genomes Phase 3 reference panels which consist of a diverse set of haplotypes from different ethnicities matching the other ethnicities in the UKBB cohort.

The imputation has been performed in groups of ~50,000 markers on 5000 samples per time. This resulted in a dataset of 92,693,895 markers in 487,442 individuals.

The mechanism of genotype imputation relies on the concept of linkage disequilibrium (LD, see section 3.2 below). LD occurs when adjacent loci are inherited together non-randomly. This could result in observing groups (blocks) of shared alleles between individuals who inherited them from the same ancestor (i.e., haplotypes). This concept facilitates the process of genotype prediction for indirectly assayed genotypes. In the imputation process, information from the directly assayed markers in the study sample is first compared to a reference panel of haplotypes which consists of information on a much larger number of markers. Then, shared haplotypes are recognised and missing data in the study sample are filled using information from the matching reference haplotype. Figure 5 below visually illustrates imputation process.

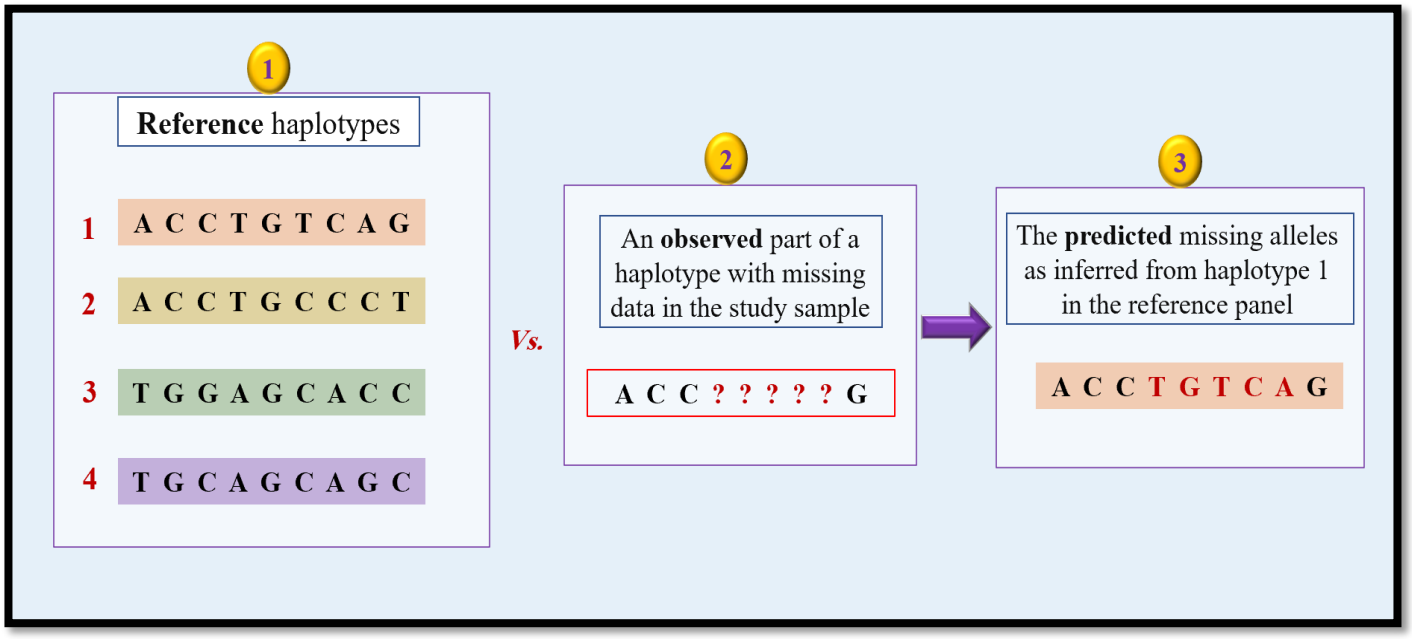


Figure 5: The imputation process and predicting missing data for the study sample.

** Using data from reference haplotypes (1), the missing data in the study sample (2) can be predicted (3) from the matching reference haplotype in the reference panel (1).*

2.1.3 Funding bodies

The UKBB study is mainly funded by the UK Medical Research Council and Wellcome Trust. Other contributors include Scottish and Welsh Governments, British Heart Foundation, North West Development Agency, and Diabetes UK.

2.2 Genetics of Diabetes Audit and Research in Tayside Scotland (GoDARTs) database

The information below has been summarized from reference 113.

2.2.1 Data collection and participants

The study started in 1996 as a collaboration between the University of Dundee's medical department, Ninewells Hospital, Perth Royal Infirmary, Stracathro Hospital and a group of Tayside general practitioners. The study focused on recruiting diabetic patients in the Tayside area. In 1998, patients were invited to provide clinical information and samples for the benefit of genetic research in diabetes mellitus (DM) and its treatments.

Data were collected in three stages in the periods 1998-2004, 2004-2009, and 2009-2015 until it reached to a total of 18306 participants (10149 type 2 diabetic patients and 8157 controls). All participants provided permission to use their medical data anonymously for research-related purposes. The GoDARTs database now consists of continually updated longitudinal prescribing and biochemistry data for the participants.

2.2.2 Genotype data

Out of 10149 cases and 8157 controls, genotype data are available for 8564 and 4586 patients, respectively. Five assay methods were used to genotype participants. Affymetrix Genome-Wide Human SNP Array 6.0 (932,979 SNPs) and the Illumina HumanOmniExpress (731,296 SNPs) were used to produce genome-wide data for 7857 cases and 1108 controls. Further customized genotyping for 707 cases and 3478 controls was performed using the Immunochip (196,524 variants, focused on immune diseases), the Cardio-Metabochip (Metabochip) (196,725 markers, focused on cardiometabolic diseases) and the Human Exome array (247,870 markers).

2.2.3 Funding bodies

The study is funded by the Scottish Home and Health Department, the Wellcome Trust, Tenovus Tayside, and the Robertson Trust.

2.3 Generation Scotland: Scottish Family Health Study (GS: SFHS) database

The information below has been summarized from references 114-116.

2.3.1 Data collection and participants

The Generation Scotland (GS) study started in 2003 as a collaboration between Scottish medical schools and the National Health Service (NHS). In 2004, GS was funded by the Scottish government to start the Scottish Family Health Study (SFHS) . SFHS collected participants as families rather than individuals to facilitate genetic epidemiological studies and uncover rare variants associated with diseases and drug response. In the period from 2006 to 2010, a total of 126,000 potential candidates were invited to participate, of whom 6665 responded and met the study

criteria (aged >= 18 – 65 and able to bring at least one family member to participate). An additional 1288 individuals participated without invitation with 16,007 family members (some participants aged above 65 (up to 100 years) and some with no relatives were allowed to participate). This gives a final of 23,960 participants in the study. 99 % of the sample were of White ethnicity, and 74.2 % shared their medication history. All participants provided consent to use their medical records for research purposes. The NHS Tayside committee on research ethics provided ethical approval for this study. The data collection started cross-sectionally with subsequent longitudinal medical data as a result of linkage to NHS data.

2.3.2 Genotype data

98% of the cohort have had their DNA extracted. Over than 20,000 participants were genotyped using high-density genome-wide chips: Illumina OmniExpress (~700,000 variants) and the human exome chip (~250,000 variants). Genotype quality checks for 32 SNPs on a test sample of ~ 10,000 subjects showed high quality genotyping with call rates of 97% or more for the majority of SNPs. The prescribing data were available for ~ 10K participants at the time of our study.

2.3.3 Funding bodies

GS was mainly funded by the Chief Scientist Office of the Scottish Government and the Scottish Funding Council.

2.4 Scottish Health Research Register (SHARE) and Genetics of SHARE (GoSHARE) database

The information below has been summarized from references 117-119.

2.4.1 Data collection and participants.

The SHARE project has established an approach to ask the public to share their data stored in NHS databases for research purposes. Any individual from Scotland aged >= 16 years can join the project. The recruitment process started in 2011 and is still ongoing with the hope of reaching 1,000,000 individuals. To date (2020), 277,082 individuals have participated in paper forms or online. The GoSHARE project, which initiated in Tayside but is now national, asks SHARE participants to consent to use of leftover blood samples collected as part of routine clinical care. From SHARE, 70,790 participants were resident in Tayside and Fife, and have linkage to medical record data; of these genotype data were available for 4,819 participants.

2.4.2 Funding bodies

The project is mainly funded by NHS Research Scotland (NRS), which is a cooperation between the Scottish Health Boards and the Scottish Government Chief Scientists Office (CSO).

2.5 Utilization of the four cohorts in our project.

The 3 Scottish cohorts have been combined in a single cohort, and then this single cohort was utilized in chapters III and IV in order to generate our initial (discovery) findings. The UKBB cross-sectional cohort was also used as a separate discovery cohort in these two chapters to generate findings using a different phenotype than the ones used in the combined Scottish cohort.

Then, in chapter V, the UKBB primary care data was used as a replication/validation cohort for the top results from the combined Scottish cohort.

Figure 6 below summarizes the process of using the four cohorts in our project.

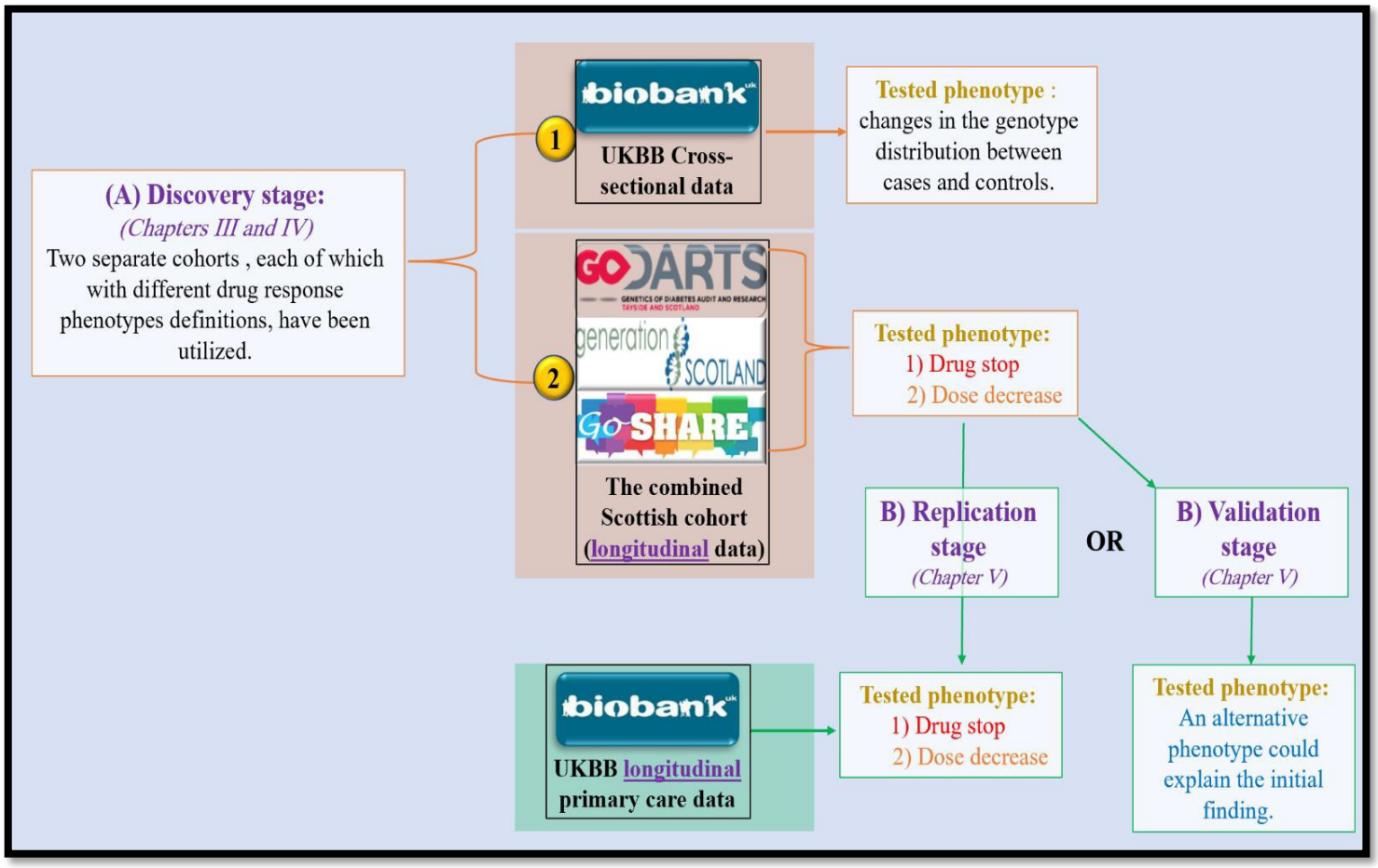


Figure 6: The process of utilizing the four cohorts (GoDARTs, GS, GoSHARE, and UKBB) in order to generate the discovery and the replication/validation results in our project.

3. Statistical Genetics

Given the role of genetics within this thesis, here I outline statistical genetic principles in general and how these were applied during the PhD.

Information in sections 3.1-3.7 below has been summarized from the reference number 120.

3.1 Alleles and genotypes

Each individual inherits a group of different genotypes. An individual's genotype consists of two alleles inherited from the parents. One is usually known as the major allele, and the other one is the minor allele based on their frequencies in a certain population. In general, the minor allele is responsible for presenting the uncommon phenotype, which is the subject of many genetic studies. The minor allele is considered common if its distribution in a specific population is over than 1% even though a frequency of >= 5% is more desirable for some researchers.

3.2 Linkage Disequilibrium (LD)

Linkage equilibrium occurs when there is a high chance for two DNA segments to separate from each other, by a process called recombination, as a result of the large distance between them. In contrast, linkage disequilibrium occurs when adjacent loci are inherited together non-randomly due to their physical proximity. LD tends to decay throughout generations.

Two statistical measures can be used to quantify the extent of occurrence of a recombination event between two variants: the absolute value of deviation from the equilibrium status (D′) and more commonly the absolute value for the correlation coefficient (r^2). If we have two linked variants with alleles Aa and Bb, then four

possible haplotypes are expected: AB, Ab, aB, and ab. D and r^2 between the two variants are given by the below equations:

$$D = AB\ frequency - A\ frequency * B\ frequency.$$

$$r^2 = D^2 / A\ frequency * a\ frequency * B\ frequency * b\ frequency$$

If the absolute value of D or $r^2 = 1$, then the two genetic variants are in complete correlation (LD) with each other and allele frequency of the two is similar. The higher the deviation from 1, the higher chance that the two variants are independent of each other.

3.3 Allele frequency calculations

Given that the major allele and the minor allele are denoted by "A" and "a" symbol respectively, three probabilities are expected to represent an individual's genotype: 'AA' (dominant homozygous), 'Aa' (heterozygous), or 'aa' (recessive homozygous). When at equilibrium, the frequency of each allele can be calculated from known genotype counts after calculating allele counts by utilizing the below equations:

*Allele A count = (genotype AA count * 2) + genotype Aa count*

*Allele a count = (genotype aa count * 2) + genotype Aa count*

Allele A frequency = Allele A count / total alleles' counts (A+a)

Allele a frequency = Allele a count / total alleles' counts (A+a)

For example, if the number of subjects with 'AA', 'Aa', and 'aa' genotypes are 1400,500, and 40 respectively, the minor allele (a) frequency (MAF) would be: -

*[(40*2) + 500] / [(1400*2) + 500] + [(40*2) + 500] = 580 / (3300 + 580) =* ***0.14***

As the total proportion of both alleles always equals 1, the major allele (A) frequency would be: *1 – 0.14 =* ***0.86.***

3.4 Genotype frequency calculations

The frequency of a specific genotype can be calculated by dividing the genotype count by the total sample size. For instance, the 'aa' and 'Aa' genotype frequencies from the aforementioned example are $40 / (1400+500+40) = 0.02$ and 0.25 respectively which implies that 'AA' genotype frequency is $1 - (0.02 + 0.25) = \mathbf{0.73}$.

Understanding the above principles is helpful in understanding what is known as Hardy-Weinberg equilibrium law and its associated calculations.

3.5 Hardy-Weinberg equilibrium (HWE) law and its assumptions

HWE law shows to what extent allele and genotype frequencies would change from one generation to another generation. It concludes that no change in the allele frequencies is expected to be seen in the next generation as long as five assumptions have been met. These assumptions are:

a) *Random selection*: carrying risky genotypes which affect survival rate would result in changing allele frequency in the next generations.
b) *Absence of new mutations:* frequency rates of alleles may change if new alleles appeared by certain mutations.
c) *No migration:* leaving one population to another population can change allele and genotype frequencies.
d) *No chance events:* occurrence of genetic drift by chance when, unexpectedly, some individuals started to contribute higher alleles than others.
e) *Random mating:* If mating occurs in a way depending on the genetic constitution between couples, genotype frequency but not allele frequency

would change. This can be seen in communities where there is a preference of mating between the same racial group for example.

A population is said not to be at HWE if any of the above assumptions has been violated.

3.6 HWE calculations and testing

HWE calculates genotype frequencies of the next generation using observed allele frequencies of the current population. Using the example in section 3.3 above, for an 'AA' genotype to be produced, both parents should provide an 'A' allele to their child. So, the expected 'AA' genotype frequency is expressed as the probability of a sperm containing an 'A' allele (0.86) multiplied by the probability that an egg provides an 'A' allele as well (0.86) = **0.7396**. Similarly, the expected 'aa' genotype frequency would be expressed as a (0.14) * a (0.14) = **0.0196**. For the 'Aa' genotype, there are two possibilities for the sperm and the egg to be fertilized, the sperm may carry either 'A' or 'a' allele, and the egg may carry either of the two alleles as well. Therefore, the expected 'Aa' genotype frequency can be calculated as follows: 2 * A (0.86) * a (0.14) = **0.2408**. Adding the 3 probabilities together (**0.7396** + **0.0196** + **0.2408**) = **1**. The net of the 3 genotype frequencies always equals 1. Accordingly, the HWE equation has been formulated to generalize the above concept.

HWE equation is expressed as follows:

$$p2 + 2*(p*q) + q2 = 1$$

Where p = 'A' and q = 'a' in our example.

The expected genotype frequencies calculated above can then be converted into the number of individuals carrying the genotypes by multiplying these frequencies by the

total sample size. 'AA', 'Aa', and 'aa' genotype counts = (1400+500+40 =**1940**) * 0.73 = **1416.2**, 1940 * 0.24 = **465.6**, and 1940 * 0.019 = **36.86**. Then, a table with the observed and expected genotype counts can be created with Chi-square test used to test whether there is a significant deviation away from HWE. The table, in our example, would look as follows:

Genotype	The observed genotype count	The expected genotype count
AA	1400	1416.2
Aa	500	465.6
aa	40	36.86

Table 2: an example for the observed number carrying the 3 different genotypes and the calculated expected number of individuals carrying these genotypes in the next generation. Any significant difference would indicate violating HWE principle.

If the p-value of the Chi-square test >= 0.05, it can be concluded that the population is at HWE. On the other hand, if the p-value is < 0.05, this indicates a deviation from HWE as there is a significant difference in the genotype frequencies between the observed population and the next generation. Visual, but not statistical, inspection of the above table would tell us that there is no significant difference between the observed and the expected genotype counts and therefore the population seems to be on HWE. It is worth noting that the HWE p-value threshold can be more relaxed with large sample sizes, as even very minor deviation away from HWE will be statistically significant but not represent clinically relevant deviation or genotyping error.

Some researchers consider p-values >= $1*10^{-4}$ or $1*10^{-8}$ as the cut-off points. In fact, this is a debatable subject among researchers.

3.7 Types of genetic models and their uses.

In genetic studies, the three genotypes can be grouped by different methods to study the genotype-phenotype association. There are four main genetic models which can be used which are explained below.

a) *Dominant model:* used when one or two copies of the variant allele (a) is/are needed for an x-fold increase/decrease in the risk linked to the phenotype. Here both 'Aa' and 'aa' genotypes are combined and compared to the reference 'AA' genotype (('Aa' + 'aa') vs 'AA').

b) *Recessive model*: used when two copies of the variant allele (a) are needed for x-fold increase/decrease in the risk linked to the phenotype. Here the 'aa' genotype is compared to the reference genotype, which represents both 'AA and 'Aa' genotypes ('aa' vs ('AA' + 'Aa')).

c) *Codominant (or genotypic) model*: used to identify the individual genotypic effect on the phenotype of interest. Here, the 'Aa' genotype alone is compared to the reference 'AA' genotype and the 'aa' genotype is compared to the reference 'AA' genotype (('Aa' vs 'AA') and ('aa' vs. 'AA')).

d) *Log-additive model*: used when one copy of the variant allele (a) is needed for an x-fold increase/decrease in the risk linked to the phenotype and two copies (aa) are required for a 2x-fold increase/decrease in the risk. The model tests whether log

ratio (cases counts/total sample size) changes linearly with 'AA', 'Aa', and 'aa' genotypes.

As the mode of effect is usually unknown when studying drug-gene associations, the log-additive model is commonly used in these studies. If the model returns a significant result, the researcher may be interested in knowing the individual genotypic effect, which is the function of the codominant model. This model shows the effect of either 'Aa' or 'aa' genotype compared to the wild-type genotype (AA).

If there are a very limited number of individuals with a recessive homozygous genotype (aa), some researchers prefer to convert the model into the dominant model where both 'Aa' and 'aa' genotypes are combined together and compared to the normal genotype. In other situations, the effect of the 'Aa' genotype would be negligible and close to the normal genotype (AA) carriers, converting the model into recessive by combing 'AA' and 'Aa' carriers in the same category would show the effect of 'aa' genotype more clearly.

3.8 Utilization of statistical genetics principles in our project

In chapter III, MAF, LD, and HWE tests have been utilized to facilitate the selection of candidate genetic variants.

In our project, we were interested in common genetic variants with MAF of 5 % or more. After calculating the number of individuals carrying each of the 3 genotypes, I then calculated MAFs for SNPs of interest using "*maf ()*" function under "*HardyWeinberg*" R package in R software and including SNPs with MAF >= 5%.

Another helpful tool to select SNPs is the LD test. Among a large list of genetic variants, many of these can be correlated to each other, and therefore it is

unnecessarily to study all of them. The LD test provides an estimation for the degree of correlation between all possible SNP pairs so it can be used to condense the initial list of SNPs into a smaller list. I performed the LD test by using *"genotype()"* and *"LD.data.frame()"* functions under the *"genetics"* package in R for all of the SNPs of interest. This helped to identify correlated SNP pairs (defined as SNP pairs with r^2 >=0.5) so we only keep uncorrelated SNPs in our analysis considering r^2 criteria.

The third useful test facilitating SNP selection is HWE. It can be used to check the validity of genotyping or the correctness of sample selection. Genetic variants which deviate from HWE would indicate that either the genotype data are unreliable due to technical error or that the sample contains different ethnic subgroups. The "*HardyWeinberg*" R package was used to calculate p-values of the HWE using "*HWChisq()"* function for SNPs of interest. SNPs were included if their p-values were >= $1*10^{-8}$. SNPs with p-values less than this threshold were considered significantly deviated from HWE and regarded as unreliable.

The HWE principle and their underlying assumptions were also utilized in order to explore drug response phenotypes in the cross-sectional databases. This will be discussed in more details later in this chapter under section 5.1.

Regarding genetic models, using *'SNPassoc'* R package', I have mainly used the log-additive model for all of our genotype-phenotype associations in chapters III, IV, and V. I have used the codominant and dominant models for some special cases in chapter V.

4. General statistics

After discussing methods related to the genotype data and before discussing our drug response phenotype definitions, in this section, I highlight methods used to study the association between the genotype and the phenotype. By knowing these methods, we can understand how I have utilized them to study the associations between genotypes (explanatory variables) and different drug response phenotypes (outcome variables) which I will discuss in section 5 following this section.

Information in sections 4.1-4.3 below has been summarized from references 121-123.

4.1 Regression model types

Regression tests examine the association between dependent (outcome/response/Y) variable and independent (explanatory/predictor/X) variable/s. There are four main types of regression models, and the type of variables on the Y- and X-axis determines the type of the regression model, which are as follows:

a) *Simple linear regression*: study the relationship between a single continuous (numerical) outcome variable and a single continuous explanatory variable such as the correlation between cholesterol level and weight.

b) *Multiple regression:* study the relationship between a single continuous (numerical) outcome variable and multiple continuous or categorical explanatory variables such as the correlation between cholesterol level and age, sex, and weight.

c) *Mixed (multilevel) regression:* study the relationship between a single continuous (numerical) outcome variable with multiple levels (e.g., multiple cholesterol measurements per patient) and multiple continuous or categorical explanatory variables such as sex and weight.

d) *Logistic regression*: study the relationship between a binary categorical outcome variable and multiple continuous or categorical explanatory variables such as studying whether or not (yes/no) hypertension is associated with age and/or diabetes mellitus.

Figure 7 below displays the four types of regression models as ordered above.

Figure 7: visual presentation of differences between simple **(a)**, multiple **(b)**, multilevel **(c)**, and logistic **(d)** regression models.

4.2 Checking regression model assumptions

4.2.1 Simple, multiple, and multilevel regression models

Simple, multiple, and multilevel regression models assume:

a) *Constant variance*: the spread of the outcome variable about its average value (i.e., residuals) is the same for all X variables (homogeneity of variance).
b) *Linearity:* the average value of the outcome variable is a linear function of the explanatory variables.
c) *Normality*: normal distribution of residuals.
d) *No multicollinearity* (i.e., no high correlation between explanatory variables): variance inflation factor (VIF) for each explanatory variable has to be < 5 to exclude multicollinearity.
e) There are *no outliers* (extreme values) in the continuous explanatory variables: an observation is said to be an outlier if the absolute value of standardized residual is >3 (> +3 or > -3).

Using statistical software such as R, normal probability plot (normal Q-Q plot) can be used to check for the normality assumption. In Figure 8 below, the right-hand plot below shows that residuals are normally distributed while the left-hand plot displays skewness in the distribution.

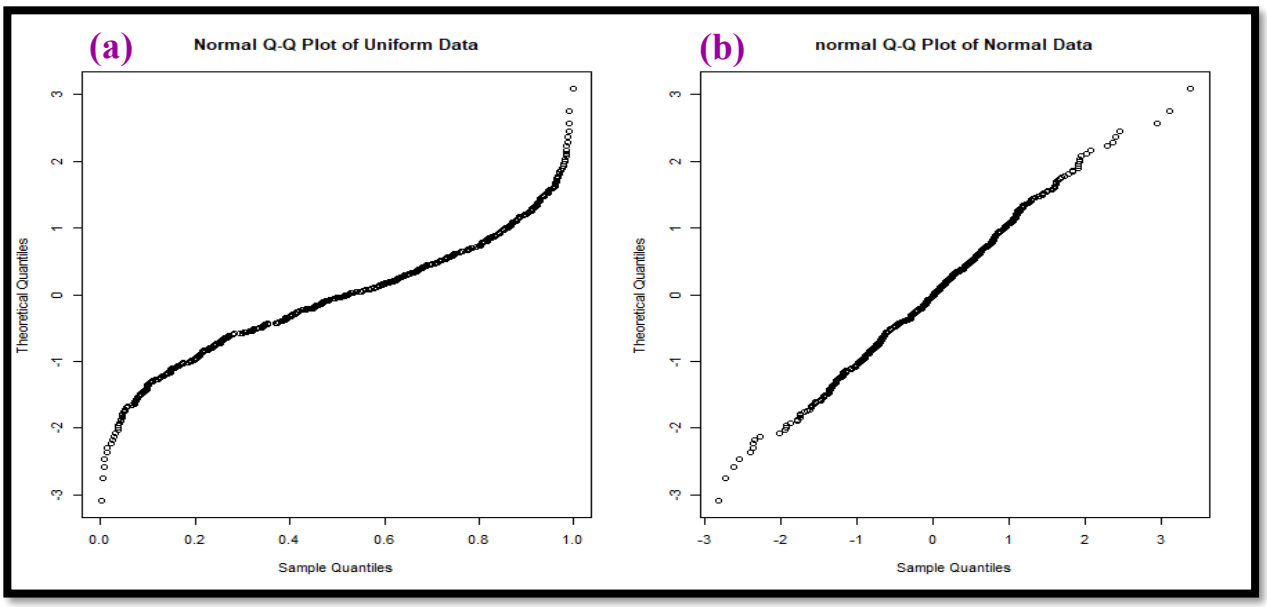


Figure 8: The difference between non-normally **(a)** and normally **(b)** distributed data when checking for the normality assumption of the regression model.

Regarding constant variance and linearity assumptions, this can be checked via residuals versus fit plots. In Figure 9 below, the right-hand plot below violates the linearity assumption since the data points are distributed in a curved line. The plot in the middle violates the constant variance assumption as the data points show a clear triangle pattern. The left-hand plot meets both constant variance and linearity assumption as the data points are equally distributed in the graph without any pattern can be observed.

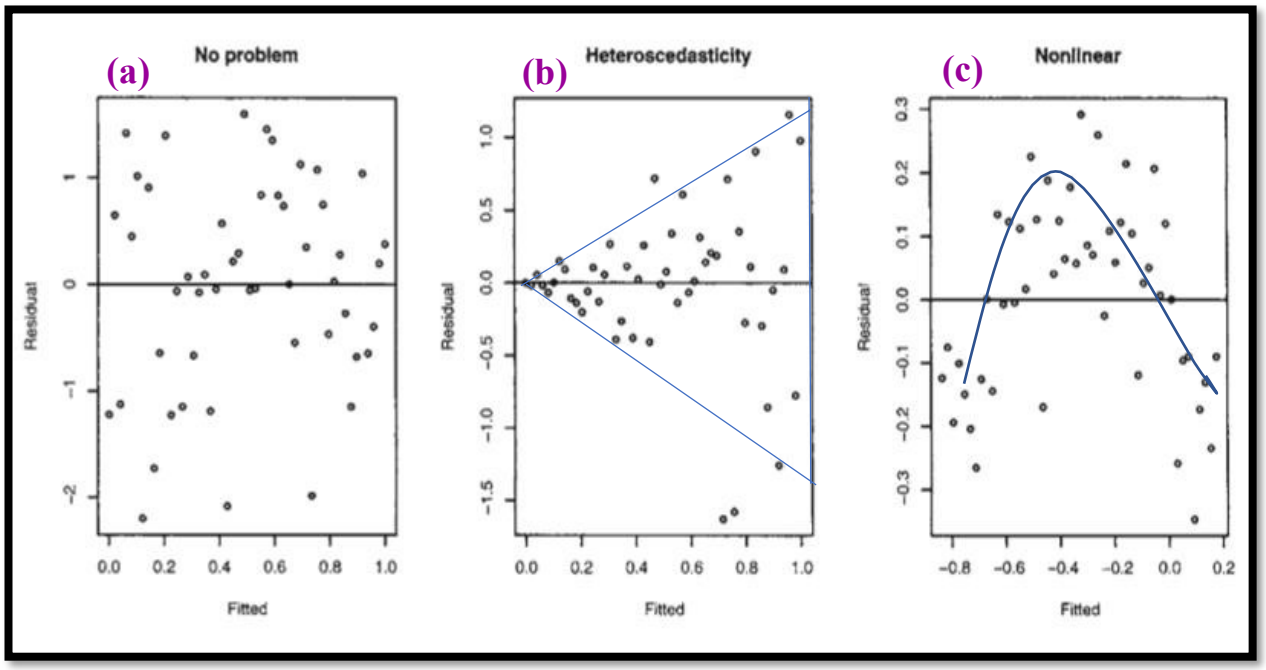


Figure 9: The differences between data meeting both constant variance and linearity assumptions **(a)** and data violating the constant variant assumption **(b)** or violating the linearity assumption **(c)** in the regression model.

4.2.2 Logistic regression

Logistic regression assumes that:

a) There is a *linear correlation* between the log of the response variable and each continuous predictor variable: can be checked the same way as above after creating scatter plots for each explanatory plot against the log of the response variable.

b) There are *no outliers* (extreme values) in the continuous explanatory variables:

an observation is said to be an outlier if the absolute value of standardized residual is >3 (> +3 or > -3). Standardized residual values are calculated by statistical software.

c) There is *no high correlation* between explanatory variables: variance inflation factor (VIF) for each explanatory variable has to be < 5 to exclude multicollinearity.

4.3 Interpretation of the main elements in regression model outputs

The simple regression is expressed in the equation:

$$Y\ variable = \alpha + \beta * X\ variable.$$

The output calculates both the intercept (α) and the slop (β) of the population regression line. This enables the prediction of Y for a known X value, such as predicting drug clearance (Y) for a patient weighing 75 Kgs (X).

The main elements of a regression model outputs are:

a) *P-value of the explanatory (X) variable*: this tests the null hypothesis that the explanatory variable is not associated with the phenotype of interest in Y-axis (H=0). For a single test analysis (e.g. 1 SNP vs. 1 phenotype), the p-value significance level is <= 0.05. In case of testing multiple hypotheses (e.g., 50 independent SNPs for 1 phenotype), the p-value significance level is modified by Bonferroni correction to be: 0.05/number of tests (e.g., 0.05/50 = 0.001 for the above example). If the p-value is below the significance threshold, then we have sufficient evidence to reject the null hypothesis that there is no association and accept the alternative hypothesis that there is an association between the studied variable and the phenotype (H≠0).

b) *β (coefficient/estimate) value*: this represents the size of the effect associated with each explanatory variable. The 95% confidence interval (CI) range [y1-y2] of the value is also usually calculated. If the β = 0.34 and 95% CI = [0.15-0.60], Y = drug clearance, X = patients' weight, p=0.003, then this could be interpreted by mentioning that we are 95% confident that average clearance of patients increases (as the β value sign is positive) by a range

between 0.15-0.60 with our best estimate of 0.34 litre/hour for every 1 kg increase in weight. This can be generalized into: we are 95% confident that average "Y" of the sample increases/decreases by "β" units for every one unit increase in "X". This interpretation applies for the simple, multiple, and multilevel regression models whilst noting that the other explanatory variables in multiple or multilevel models are constant. In the logistic regression model, the β value represents the odds ratio (OR) which is usually calculated with its CI. The OR can also be converted into a percentage to deliver another way of interpretation by using the absolute result of the equation: (OR * 100)-100. If the explanatory variable is continuous, then this could be interpreted as one unit increase in the explanatory variable is expected to increase/decrease the odds of the risk factor in the Y-axis by OR times or percentage. If the explanatory variable is categorical, then one or more of the categories would be OR times more/less likely to experience the risk factor in the research question.

c) *R-square (coefficient of determination)*: this value shows the percentage of the variability in "Y" which can be explained by variation in "X". If R-square = 90% for the above example; this implies that 90% of the variability in clearance is explained by the variability in weight. As R-square increases with adding multiple variables in the model, the adjusted R-square value is used within multiple regression models rather than the raw value.

4.4 Utilization of regression models in this thesis.

In this thesis, particularly in chapters III and IV, I have utilized the logistic regression model for most analyses since our main outcome variables were categorical (On the drug/drug combination (Yes/No), drug-stop (Yes/No), dose-decrease (Yes/No)). In chapter V, I have used a continuous outcome variable (systolic blood pressure (SBP)) in two different ways: one time as a single measurement per patient (e.g., mean SBP) and once as two measurements per patient (e.g., mean SBP pre-and post-treatment). Therefore, a multiple regression model was used for the first case, and the mixed effect model was used in the second case. *"Lm ()"* and "*glm ()"* functions in the *"stats"* package in R are usually used for multiple and logistic regression analysis respectively. However, I have utilized more sophisticated functions (*"WGassociation()"* and *"association()"*) under *"SNPassoc*" R package, which are specially designed for genetic analysis. The analysis for both categorical (e.g., cases/controls) or quantitative (e.g., SBP) outcome variables can be run using these functions to provide the results under any genetic model of interest (which have been explained previously in sub-section 3.7 of this chapter). Beside p-values, the results given are "ORs" (for categorical outcome variables) or "mean difference" (for quantitative outcome variables) with their 95% CI. Regarding the multilevel (mixed) regression model, it can be fitted using *"Lme()" or "Lmer()"* functions under *"nlme"* or *"lme4"* packages respectively. While these packages do not compute p-values, *"Lmer()"* function under *"LmerTest"* package provides p-values in the model and is the package used in chapter IV.

5. Defining drug response phenotypes.

5.1 In a cross-sectional cohort (the UKBB)

I started this PhD project in March 2017 when the prescribing data for the UKBB cohort were cross-sectional. As certain drug response phenotypes such as dose change, drug-stop, or changes in clinical measurements cannot be defined using prescribing data collected from patients once and without a follow up, defining drug response or intolerance is not directly possible. However, it is possible to infer a tolerant/efficacy phenotype using the information about genotype distribution after applying assumptions underlining the HWE principle (review section 3.5). For instance, if the main interest was studying the association between a specific single nucleotide polymorphism (SNP) and statin-induced myopathy in the absence of longitudinal prescribing data, clinical measurements (e.g., creatine kinase), and GP reports about this side effect, the genotype distribution can be used to detect deviation away from HWE. According to the HWE principle, the random selection of any two groups within the same population will show no difference between them in the allele and genotype frequency. However, if the presence of a particular allele causes a person to stop the statin, then that allele will be underrepresented amongst statin users compared to the unselected population, i.e. there will be deviation away from HWE. Similarly, if that allele is associated with good response to statins, the allele will be over-represented in the statin-treated group. Therefore, HWE calculations should only performed on controls. In controls, there is no identifiable risk factor which could affect the distribution of the variant allele and, therefore, any deviation from HWE could be more likely explained by the presence of genotyping errors or different ethnicities. In contrast, deviations from HWE in cases, could be more likely explained by the presence of the risk factor rather than genotyping errors.

The drug response phenotypes recognized by this method would be classified into two categories, as follows:

a) *Increased therapeutic efficacy phenotype*: this occurs when the variant allele found to be highly distributed within the drug users compared to non-users.
b) *Drug resistance/switch or increased toxicity phenotype*: this occurs when the variant allele found to be significantly lower within the drug users compared to non-users.

Figure 10 below displays different scenarios of drug response phenotypes based on genotype distribution compared to the distribution in the normal situation.

However, increased or decreased the distribution of the variant allele among drug users are not always a reflection of changes in drug response as these might also reflect an association with diseases for which the drug is prescribed for. For example, if a genetic variant was observed overrepresented among statin users, this wouldn't necessarily imply that this variant is associated with increased statin response; it may also imply that this variant could be associated with increased risk of hyperlipidemia. To deal with this issue, we investigate significant hits from the UKBB for their associations with diseases (this will be explained later in section 6.3 of this chapter) and also, we examine the same drugs using different and more accurate phenotypes identified from the longitudinal prescribing data from the combined cohort which is explained in the next section.

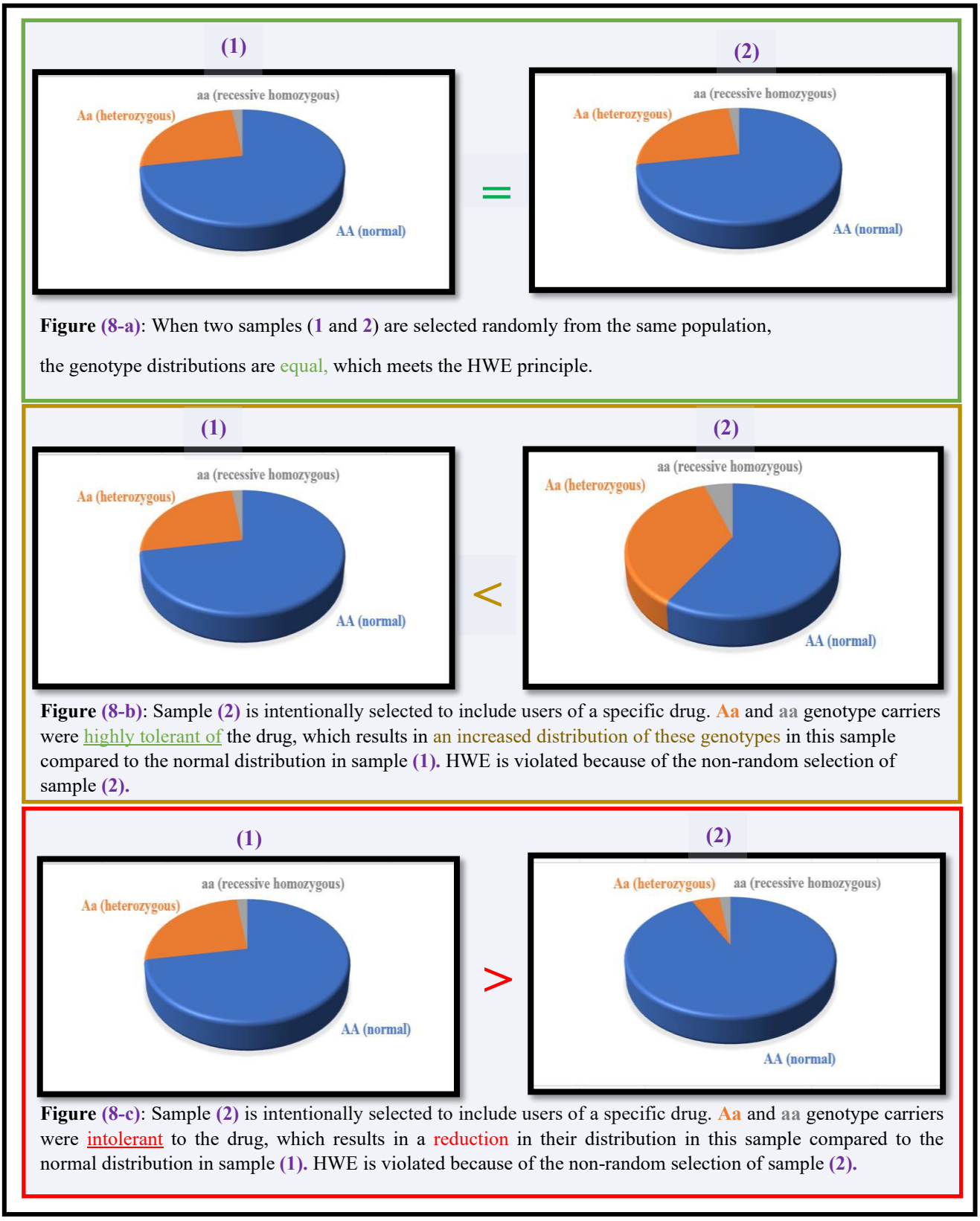


Figure (8-a): When two samples (1 and 2) are selected randomly from the same population, the genotype distributions are equal, which meets the HWE principle.

Figure (8-b): Sample (2) is intentionally selected to include users of a specific drug. Aa and aa genotype carriers were highly tolerant of the drug, which results in an increased distribution of these genotypes in this sample compared to the normal distribution in sample (1). HWE is violated because of the non-random selection of sample (2).

Figure (8-c): Sample (2) is intentionally selected to include users of a specific drug. Aa and aa genotype carriers were intolerant to the drug, which results in a reduction in their distribution in this sample compared to the normal distribution in sample (1). HWE is violated because of the non-random selection of sample (2).

Figure 10: The differences in the genotype distributions between two random samples (8-a) (the normal situation) and two samples one of which represents well tolerance (8-b) or intolerance (8-c) to a specific drug.

5.2 In longitudinal cohorts (GoDARTs, GS, and GoSHARE).

There are different methods to define drug response phenotypes given the availability of follow-up prescribing data. Depending on the nature of the research, phenotypes can be either general or specific to the drug of interest.

We were interested in studying drug-gene and drug-drug-gene interactions of most commonly used drugs in the UK. As we aimed to investigate all common drugs but not only a specific drug class, we needed a method to define drug response which could be generalized to all different drugs. The most suitable phenotypes in this situation were a “drug-stop” or “dose-decrease” phenotype. If statistically significant findings emerge from this approach, then a more targeted drug-specific approach was used, for example, anti-hypertensive efficacy for blood pressure drugs. General (“drug-stop” and “dose-decrease”) and drug-specific response phenotypes for both drug-gene and drug-drug-gene interaction studies are illustrated below.

5.2.1 General drug response phenotypes: drug-gene interactions study.

Stopping the drug would indicate experiencing intolerable side effects, lack of therapeutic efficacy, or both. Decreasing the dose may be a sign of toxicity (necessitating dose reduction) or marked efficacy (potentially causing symptoms or signs of overtreatment).

For the ‘drug-stop’ phenotype, cases were those who only had one prescription of a drug that is usually prescribed chronically; controls were those who received two or more prescriptions of the same drug since prescribed (tolerability indicator). For all drugs examined, we observed that the majority of controls per drug receive 3 or more prescriptions of the same drug after the first prescription.

In general, side effects which could require dose reduction could occur any time after treatment initiation. They may occur directly, a couple of months after, or even one or more years after starting the treatment. Therefore, for "dose-decrease" phenotype any patient is considered a case when a reduction in his/her daily dose is observed at a certain point of time during the treatment period. Patients on the same treatment who have never reduced their daily dose since initiation are considered to be the control group.

5.2.2 General drug response phenotypes: drug-drug-gene interactions study.

A similar approach to defining drug-stop and dose-decrease phenotypes was applied to drug-drug interaction studies. After identifying patients who have concurrently prescribed the drug combinations of interest, they are classified into cases and controls. Those who have been prescribed drug 1 or drug 2, which are both known to be used chronically, only once and this single prescription has been prescribed during the period of taking the other drug (the interaction time) are considered cases. The control group represents all other patients who are on the same drug combination but have never stopped any of the two drugs during the interaction time (as indicated by two or more prescriptions-for all combinations, the majority of controls we observed had 3 or more prescriptions of either of the two drugs during the interaction time). Similarly, for the dose-decrease phenotype, individuals who have reduced the daily dose of any of the two drugs during the interaction time are considered cases while those who do not reduce the daily dose of any of the two drugs during the period, they are co-prescribed are considered the control group.

In the clinical sitting, the undesirable drug-drug interaction outcomes can be dealt with by either stopping or reducing the dose of either the perpetrator or the victim drug. If the interaction is known, the decision regarding which drug to stop or to reduce its dose depends on the individual patient's needs. In some cases, the undesirable outcome might have happened due to unknown DDI, and in this situation, the prescriber is more likely to make modifications to the last added drug. For either of the above scenarios, modifications can occur for drug 1 or drug 2, and this is why we have considered our case groups to represent those who have experienced the phenotype of interest considering both drugs.

Figures 11 and 12 below illustrate methods of defining drug-stop and dose-decrease phenotypes.

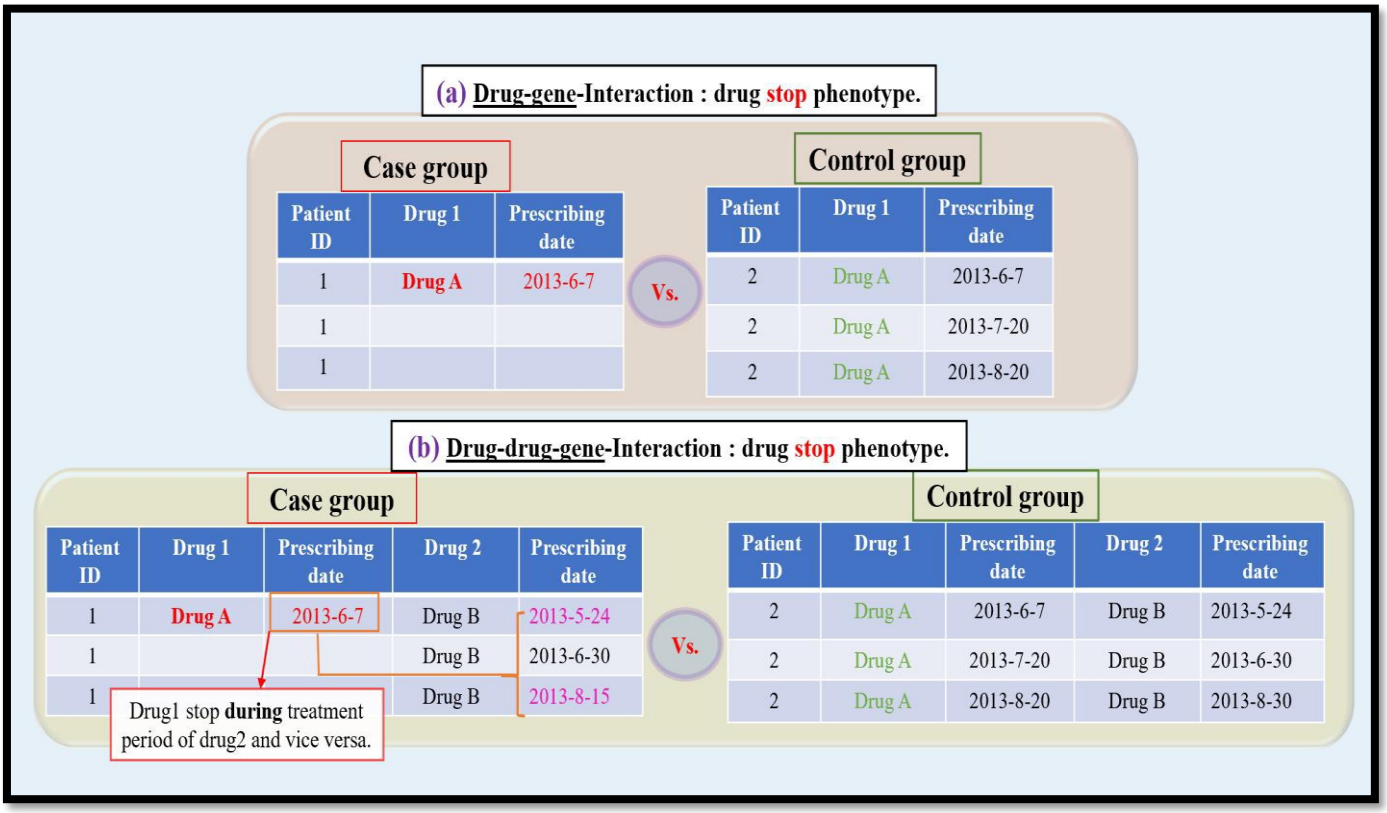


Figure 11: Methods for defining drug-stop phenotype for drug-gene (**a**) and drug-drug-gene (**b**) interactions studies. Cases (who stop drugs) are those prescribed a chronic drug once while controls are those who have been prescribed the same drug multiple times. In the drug-drug-gene interaction study, stopping the drug occurs during the interaction time with the other drug. Cases include those who stop drug 1 during treatment with drug 2 and those who stop drug 2 during treatment with drug 1.

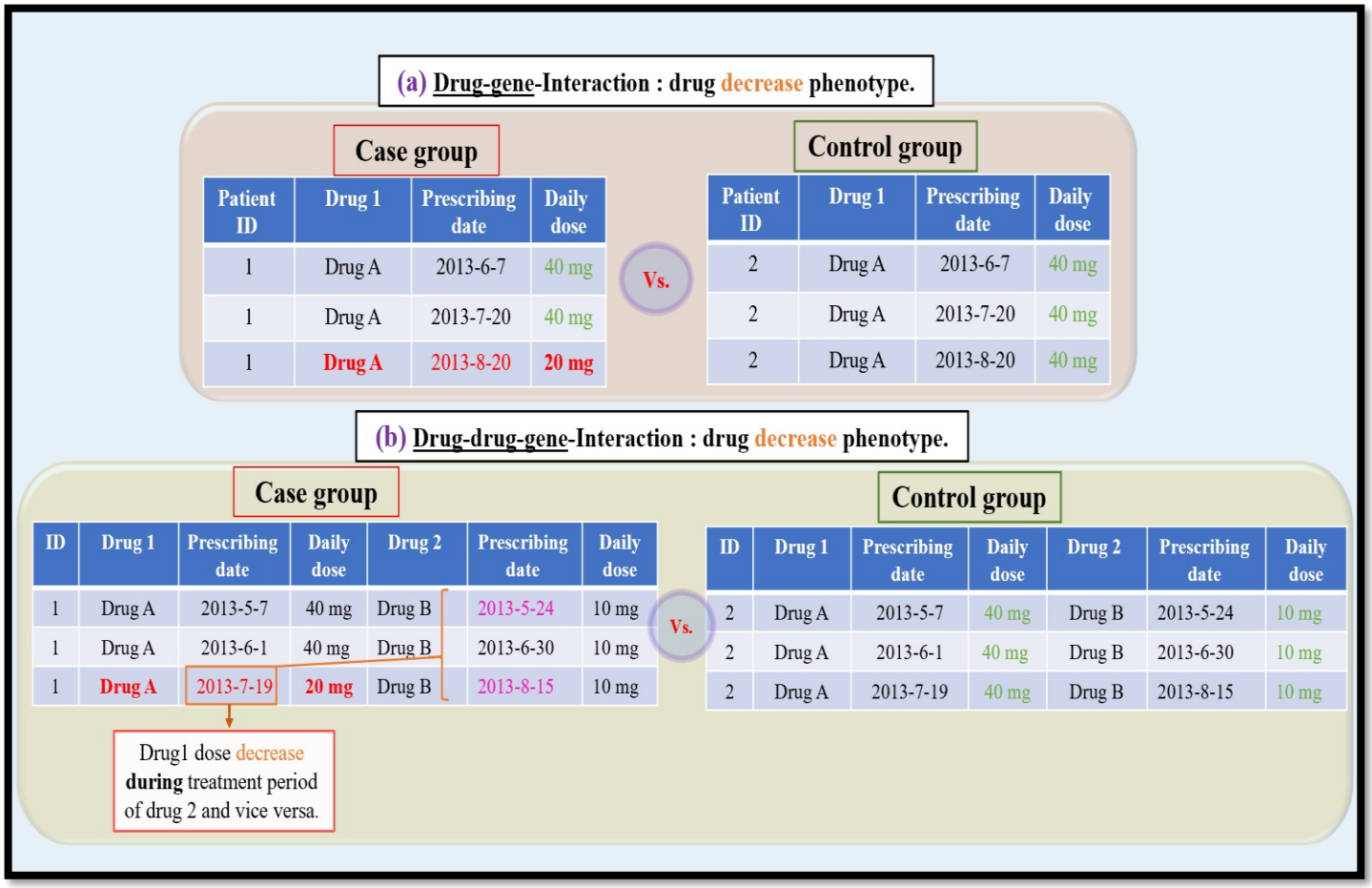


Figure 12: Methods for defining the dose-decrease phenotype for drug-gene (a) and drug-drug-gene (b) interaction studies. Cases are those who have decreased their daily dose at a certain point of time while controls are those who have never decreased the dose. In the drug-drug-gene interaction study, dose reduction should occur during the interaction time with the other drug. Cases include those who decrease drug 1 during treatment with drug 2 and those who decrease drug 2 during treatment with drug 1.

Having identified cases and controls as per the above definitions, a case-control genetic association analysis can be then run to test the association between the SNPs of interest and the drug/drug combination of interest considering the two drug response phenotypes per drug/drug combination.

5.2.3 Defining SBP response: drug-gene interactions study.

Figure 13 below demonstrates the model fitting for studying SBP response in drug-gene interaction studies.

Having identified potential drug-gene interactions in the longitudinal combined Scottish cohort, the arrival of a larger longitudinal primary care data for UK Biobank late in my PhD studies, enable me to validate signals of interest. The phenotype of interest was systolic blood pressure reduction in response to antihypertensive agents. I included white-British individuals where systolic blood pressure measurements are available for them one year before and year after starting the treatment of interest. This one-year window specification reduces the effect of other potential anti-hypertensive agents prescribed for the same patient. Then, mean systolic blood pressure is calculated for each patient for all measures in the pre-treatment, and again in the post-treatment year. A multiple regression model was then used to study the influence of the genetic variant of interest on the blood pressure response to the drug. The outcome (Y) continuous variable was the mean systolic blood pressure one-year post-treatment. The explanatory variables were the mean systolic blood pressure one-year pre-treatment (to adjust for the drug-only effect) and the SNP of interest (categorical genotypes) adjusted by age and sex.

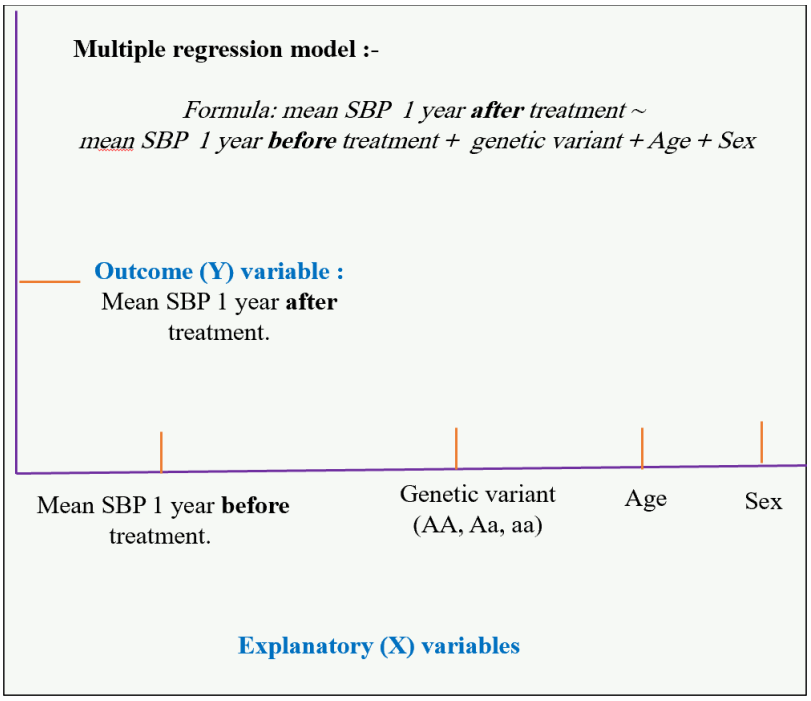


Figure 13: Visual representation of outcome and explanatory variables included in the multiple regression model to study the effect of the genetic variant on the blood pressure reduction of antihypertensive agents.

5.2.4 Defining SBP response: drug-drug-gene interactions study.

Figure 14 below illustrates steps I followed to study blood pressure changes following a certain drug-drug-gene interaction.

I employed two approaches to investigate drug-drug-gene interactions for systolic blood pressure (SBP) reduction. In the first approach, I was interested in what happened to the SBP in those treated with the antihypertensive of interest (victim drug) who started an interacting (perpetrator) drug. Initially, all patients on the drug combination of interest were identified. From these only those who started the anti-hypertensive drug first who subsequently added the perpetrator drug were included. This ensured that a stable on-treatment SBP could be defined before initiation of the perpetrator drug. Even though this model resulted in a large reduction in sample size, the benefits are of a clean within-person response to the addition of the perpetrator drug. The mean SBP 1 year before starting the antihypertensive drug, the mean SBP in the first year after initiation of antihypertensive drug alone and before adding the perpetrator drug, and the mean SBP after adding the perpetrator drug to the antihypertensive drug were then calculated. As two SBP measurements are being compared from the same patient (i.e., before and after adding the perpetrator agent), these are considered dependent measurements which violate the assumption of independence in a regression model. Therefore, a linear mixed-effect model was used in this case. The outcome (Y) variable was mean SBP after the antihypertensive drug (taken as two measurements per patient within 1 year: the first one is while on the victim (antihypertensive) drug only and the second one is after adding the perpetrator agent to the victim drug). The explanatory (X) variables were mean SBP one year before the antihypertensive drug (to adjust for the effect of SBP levels before starting the antihypertensive), the treatment context (whether the mean SBP is during the

antihypertensive drug alone or after adding the perpetrator drug), age, and sex. The resulting significance of the "*treatment context*" variable will show if there is a significant difference in SBP levels before and after adding the perpetrator agent to the antihypertensive drug.

Having examined the changes in SBP from measurements taken from the same patient, a second model was then developed for the purpose of increasing the sample size. In this model, two groups are defined – those treated with the antihypertensive alone, and those treated with the combination of antihypertensive and perpetrator drug. For the first group, all patients on the victim drug alone (i.e., have no preparator drug) are included and mean SBP 1 year after treatment is calculated. For the second group, all patients who started the perpetrator drug first and then added the victim (i.e., antihypertensive) drug are included. Mean SBP is calculated for these patients during the interaction time (i.e., after adding the antihypertensive drug to the perpetrator drug and within 1 year of interaction). Then I compare the SBP changes between these two groups. Although there is some inaccuracy of this phenotype as measurements are taken from different patients, we benefit from the advantage of greatly increasing our sample size as the first group includes all patients on the antihypertensive drug rather than only those who started the antihypertensive agent before adding the perpetrator agent as in our first model. We benefit from the large sample size in order to see if we can replicate the results from the first model and most importantly, to have enough sample for the drug-drug-gene interaction study. Our response and explanatory variables are the same as of the first model, but the difference is that in the first model, I compare SBP changes from the same subject while in the second model I compare SBP changes between two groups of patients.

As in the second model SBP measurements are taken from different subjects, these measurements are considered independent values. Therefore, a normal multiple regression model can be used here. The outcome (Y) continuous variable was mean systolic blood pressure after the victim drug (for two groups of patients, as explained above). The explanatory variables are mean systolic blood pressure 1 year before the antihypertensive drug (to adjust for both SBP levels before starting the antihypertensive drug for the first group and also to adjust for SBP levels just before the interaction time for the second group), a categorical variable answering the question whether the mean SBP belongs to group1(on the victim drug only) or group 2 (on the combination) patients, the SNP of interest as a 3-level categorical variable, age and sex. The model is first run without the genotype data to identify the effect of the drug-drug interaction on SBP; then genotype is added in to study its effect on changing SBP levels between the two groups.

Assessing compliance to the treatment in both cross-sectional and longitudinal prescribing data is quite challenging. In the cross-sectional data, there is no way to recognise this as no follow-up prescribing data are available. In the longitudinal prescribing data, compliance could be assessed via observing whether or not the drug is regularly prescribed for the same patient. However, this is a weak indicator for real compliance to the treatment as simply the drug can be dispensed but the patient is not actually taking his/her treatment regularly. In fact, non-compliance to long-term therapies is a very common issue in clinical practise. On the other hand, if the prescribing data shows irregular prescribing of a specific drug, this doesn't necessarily imply non-adherence as, for example, patients could move out of the area. Overall, we might be able to only define a measure of "coverage" but a real-world evaluation of compliance cannot be simply obtained from our data.

Step 1: Examining whether there is a significant drug-drug interaction influencing blood pressure levels occurring after adding the perpetrator drug to the victim drug for the same patient.

Step 2: Further investigation for step 1 results with a different model by comparing blood pressure changes between 2 groups: those on the victim drug alone and those on the perpetrator drug + the victim drug.

Step 3: Studying the influence of the genotype of interest on SBP changes between the two groups.

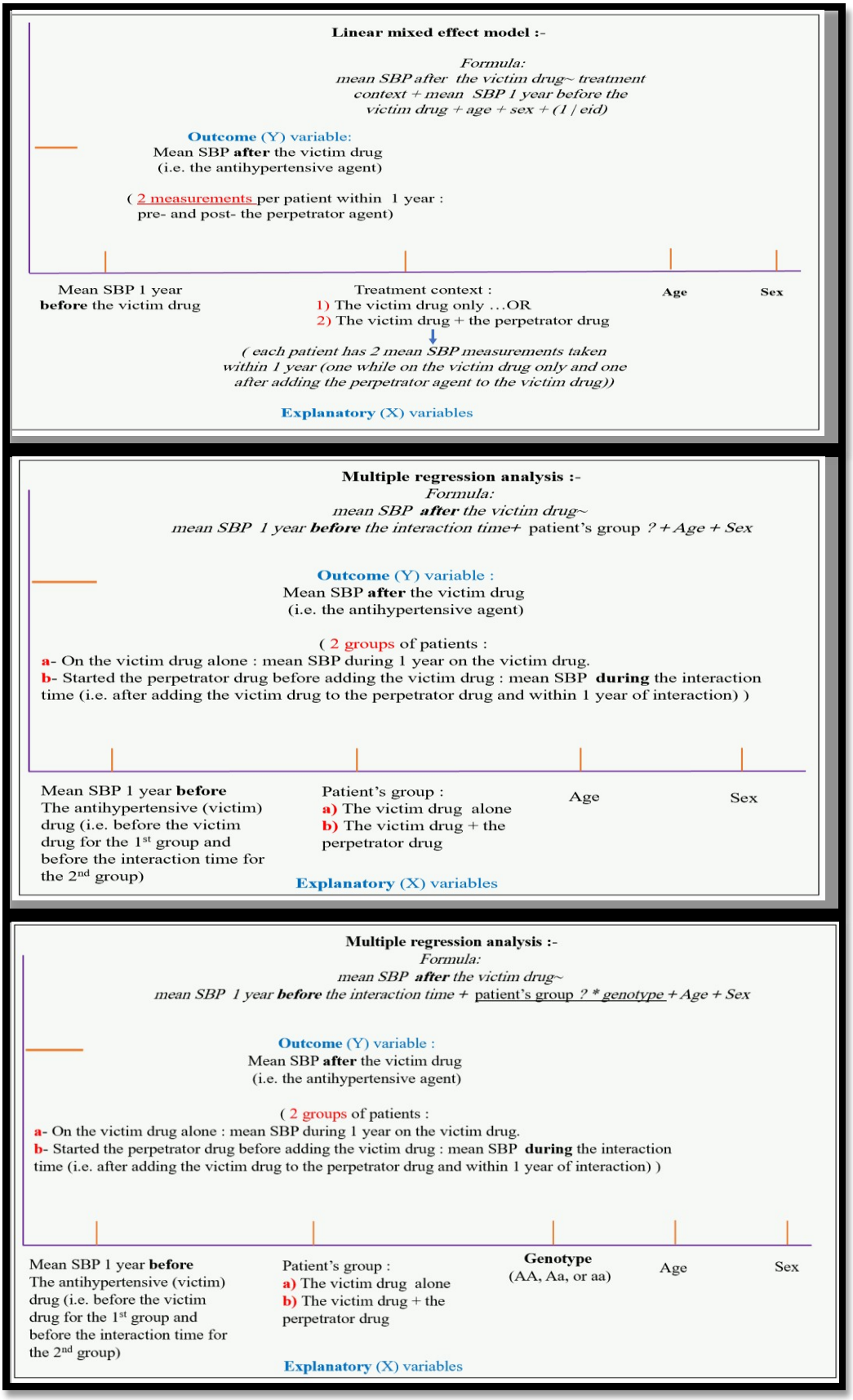


Figure 14: Steps of studying blood pressure changes following a certain drug-drug-gene interaction.

5.3 Utilization of our drug response phenotypes in our project.

In chapters III and IV, I have mainly used genotype distribution changes (UK Biobank cross-sectional data), and drug-stop and dose-decrease phenotypes (Scottish cohorts) and studied their associations with a selected group of genetic variants (focused on pharmacokinetic variants). In chapter V, beside drug-stop and dose-decrease phenotypes, I have used the change in Systolic Blood Pressure (SBP) phenotype (UK Biobank primary care data) for further investigation for some of our findings from chapters III and IV.

6. Auxiliary databases

These are the databases which were used as helpful tools in this project. Although the main findings in this thesis are not directly dependent on these resources, they were used to identify variables to be incorporated into the models or used in validation. Here I describe three auxiliary databases utilized in our work with clarifying how/where they have been used.

6.1 DrugBank database

The database description below has been summarized from reference 124.

DrugBank (https://www.drugbank.ca/) is a huge comprehensive database on drugs and is one of the most widely used worldwide. It started in 2006 by presenting chemical and physical properties of drugs approved by the FDA. The second release in 2008 included pharmacological and some pharmacogenomic data. In 2010, pharmacokinetic data with drug classification into substrates, inhibitors, or inducers for metabolizing enzymes or transporters were added. The fourth release

in 2014 added a large amount of data on drug metabolism and structure-activity relationship. The final major update in 2018 added 522 newly FDA-approved drugs with dramatic improvements on the information available in all areas mentioned above with other new enhancements. The resource is well referenced with hundreds of published papers supporting the information available.

The main purpose of utilizing DrugBank in this project was in chapter IV as it enabled us to identify potential routes of PK interaction for the common drug combinations in the UK. This can be achieved using the information on classifying drugs into substrates, inducers, or inhibitors. I used DrugBank at the beginning of this project in 2017 at the time when the last update about pharmacokinetic drug interactions information was released in 2014 (the 4th release) before the last update in 2018. In 2014, the number of identified metabolizing enzymes and transporters was 253 before it doubled in the last update into 497 reflecting that many novel drug-interactions were not recognized, and future research is likely to uncover even more.

I have also used the database for another purpose in chapters III and IV. It helped to identify whether the gene, in which a genetic variant connected with the use of a certain drug was detected, belongs to a metabolizing enzyme or a transporter related to the drug of interest (i.e., the drug is a substrate for the gene product).

6.2 Pharmacogenomic knowledge database (PharmGKB)

The database description below has been summarized from references 125 and 126.

PharmGKB (https://www.pharmgkb.org/) is a large well-known dataset specializing in pharmacogenomic studies and their clinical impacts. It started in 2000 and is continuously updated as new research emerges. The dataset contains over more than 600 drugs and describes more than 100 pharmacokinetic or pharmacodynamic pathways related to variability in drug response.

The drug-gene associations are classified into 4 categories in PharmGKB as follows:

Level 1A evidence:

Represents associations which have been already translated into clinical practice or outlined by the Clinical Pharmacogenetics Implementation Consortium (CPIC).

Level 1B evidence:

Shows associations which have been replicated multiple times in different cohorts.

Level 2A evidence:

This category is related to associations found with very important genetic variants which are more likely to be functional variants.

Level 2B evidence:

Here are associations which have conflicting results or with small effect size.

Level 3 evidence:

Under this level are significant findings from a single paper which are not replicated.

Level 4 evidence:

The association is considered weak when it is from a case report or an in vitro study only.

The main purpose of PharmGKB in this project was in chapter III to select the candidate genetic variants. As shown above, the resource provides an informative collection of genetic variants which have been previously linked to at least a single drug response phenotype, or where there are no publications specifically related to drug response but the variant could be a potential variant affecting drug response based upon in-vitro or in-silico models. I extracted all of the potential variants related to PK pathways from this database.

The other purpose of using PharmGKB was also in chapter III and IV when it helped us to identify whether any of our significant drug-gene associations had been identified previously, as recorded in the database.

6.3 GeneATLAS database

The database description below has been summarized from reference 127.

The GeneATLAS dataset (http://geneatlas.roslin.ed.ac.uk/) is a rich resource of trait-variants associations. It is a result of a study conducted by Edinburgh University research team who used a cohort of 452,264 European participants from the UKBB study and performed genome-wide association study (GWAS) to test the association between 778 different traits and over more than 30 million variants (9,113,133 after quality control filtration). Each set of traits, non-binary (n= 118) and binary (n = 660) were tested against the 9,113,133 variants with a genome-wide significance level of $< 1x10^{-8}$ for all traits.

In chapters III and IV, I used GeneATLAS as an investigational tool for our findings from the UKBB cross-sectional data. As discussed earlier, comparing those on the drug/drug combination (cases) with those who are not (controls) could yield significant results that reflect the indication for the drug (i.e., disease risk) rather than drug intolerance or efficacy. Therefore, I used GeneATLAS to investigate whether or not the significant SNP detected was linked with any of the disease phenotypes rather than the drug response phenotype I have tested.

Chapter III:

Drug-Gene Interactions for the Most Commonly Used Chronic Drugs in the UK

Abstract

There has been increasing research in pharmacokinetic drug-gene interactions over the last few years. However, our knowledge in this area remains limited, given the large number of drugs entering the market annually and the multiplicity of elimination pathways for each drug. In the present study, we attempted to discover novel and clinically important drug-gene interactions among commonly used chronic drugs in the UK. We studied the associations between 50 common chronic drugs and 162 selected genetic variants in important enzymes and transporters. We utilised two cohorts: a combined cohort of longitudinal prescribing data for three Scottish cohorts and the UK Biobank cross-sectional prescribing data. In the combined longitudinal cohorts, the drug-variant combinations were studied in relation to drug-stop or dose-decrease phenotypes while in the UK Biobank cross-sectional data, we compared the genotype distribution between drug users and non-users.

8 novel drug-gene variants were identified in the combined longitudinal cohorts, and 4 novel drug-gene variants were identified in UK Biobank. From the combined cohort, users of ramipril or metformin who carry the rs1135840 (G>C) *CYP2D6* or rs1045642 (A>G) *ABCB1* variants respectively were less likely to stop their treatment per allele: (OR (95%CI)) Ramipril/CYP2D6 0.7 (0.6-0.82), $p=1.01 \times 10^{-5}$; Metformin/ABCB1 0.75 (0.64-0.87), $p=1.636 \times 10^{-4}$. On the other hand, those with the C allele at rs152023 variant in ABCC1 were more likely to stop nifedipine per allele OR 1.32 (1.15-1.53), $p=1.168 \times 10^{-4}$. Those with the A allele at rs5788 variant in PTGS1 were more likely to stop nicorandil treatment per allele OR 1.93 (1.39-2.69), $p= 1.718 \times 10^{-4}$.

Regarding the dose-decrease phenotype, carriage of the rs4918758 (T>C) *CYP2C9* , rs9895420 (T>A) *ABCC3*, or rs868853 (C(*minor allele*)>T) *ABCC4* minor alleles have been observed to be associated with decreased likelihood to reduce the dose of quinine OR 0.71 (0.6-0.85),p=8×10^{-5}; doxazosin OR 0.54 (0.38-0.76), p= 1.2×10^{-4}; amlodipine OR 0.55 (0.4-0.75), p= 2.8×10^{-4} respectively. The rs7916649 (T>A) *CYP2C19* mutation was found connected with higher odds to decrease valproic acid dose (OR = 1.95 (1.37-2.76), p=1.475×10^{-4}).

The results from the UKBB cohort showed a decreased frequency of the minor alleles of rs555754 (G>A) *SLC22A3*, rs3743527 (C>T) *ABCC1*, or rs8187843 (G>A) *ABCC1* variants among users of lansoprazole OR 0.95 (0.93-0.97),p= 1.58×10^{-5}; bendroflumethiazide OR 0.95 (0.93-0.98),p= 4.87×10^{-5}; or gabapentin OR 0.78 (0.68-0.89),p= 2.12×10^{-4} respectively. Similarly, the minor allele frequency of rs2231135 (A>G) in ABCG2 was increased in rosuvastatin users (OR = 1.22 (1.1- 1.35), p= 3.02×10^{-4}).

Finally, in an important validation of the method used, a number of previously reported drug-gene interactions were replicated. We also show other novel and potentially important associations.

1. Introduction

Adverse effects or loss of therapeutic efficacy are commonly seen after the initiation of new drug treatment. Many factors contribute to the risk of these adverse outcomes, including age, sex, weight, and drug-drug interactions (DDIs).
DDIs can occur by two different processes: Pharmacodynamic (PD) or Pharmacokinetic (PK). PD interactions are mostly predictable, and therefore can be highly avoidable, as they include interactions between drugs with similar or opposing pharmacological effects resulting in increased toxicity or treatment failure, respectively. PK interactions, however, can occur between any combination of medications regardless of their pharmacological action, which make them more challenging to predict and avoid. These interactions occur when the precipitant drug influences the concentration of the object drug via affecting its absorption, distribution, metabolism, or excretion by inhibiting/inducing its metabolizing enzymes or transporters. In chapter I, I have shown the distribution of important transporters in different tissue types. PK interactions can also occur when patients, instead of being co-treated with the precipitant drug, carry reduced or increased activity genetic variants in these metabolizing enzymes or transporters.

There has been an increasing number of drug-gene interaction studies in recent years. The recent update (2020) from Pharmacogenomic Knowledge Base (PharmGKB) (described in chapter II) shows ~ 23,000 published studies with 4,326 genotype-based clinical annotations presenting genotype-related differences in drug response (i.e. efficacy and toxicity). In addition,139 clinical guidelines for some drugs related to dosing recommendations are now available in the database. Of note, to date, there have been 321 drugs receiving pharmacogenomic labels from the Food and

Drug Administration (FDA) stating whether testing for a specific genetic variant is required, recommended, or not necessary before drug initiation.

In spite of these recent advances, the knowledge on PK drug-gene interactions is still limited. Given the large number of drugs with new drugs available regularly and the presence of multiple elimination pathways for a single drug, it is nearly impossible for us to have DGIs studies performed on every single drug covering all of its different elimination pathways. Therefore, currently used drug-interaction tools mainly report "potential" rather than actual interactions based on clinical studies. In addition, some important interactions could be totally ignored in pharmacogenomic studies as the entire routes of drugs' absorption and elimination are not fully uncovered giving the fact that new drug targets are appearing from time to time. Thus, most of the current work on DGIs focus only on well-known drug absorption/elimination pathways and genetic variants.

In this study, I aimed to uncover novel clinically important interactions between a large variety of commonly prescribed chronic medications and selected group of genetic variants in important metabolizing enzymes and transporters by utilizing four different UK cohorts. I begin by a brief description of the cohorts before discussing the methodology and results.

2. Study populations

Four cohorts have been utilized in this study which are: the UK Biobank (UKBB), Genetics of Diabetes Audit and Research in Tayside Scotland (GoDARTs), Generation Scotland: Scottish Family Health Study (GS:SFHS), and Genetics of the Scottish Health Research Register (GoSHARE). All of these cohorts have been prescribed in chapter II. In this chapter, the Scottish longitudinal cohorts (with ~ 27K participants) and the UKBB cross-sectional cohort (with ~ 500K participants) will be utilized to generate the discovery findings for our drug-gene interactions study.

3. Methodology

The methods are summarized in the five steps below:

3.1 Selection of commonly used drugs for inclusion

The cross-sectional prescribing data from the UKBB study was utilized to identify the most frequently prescribed drugs for around 500,000 patients. More than 3100 drugs were observed as prescribed across the UK Biobank at least once. 48 drugs had duplicates, each of which has been recorded in different writing formats and were assigned different UKBB codes. A single format and UKBB code were selected, and the total drug frequency considering all of the different formats was calculated. Defining the common drugs as the drugs used by no less than a thousand participants, I have produced a list of the top 122 most frequently used drugs in the UK. From this list, we have selected the top common 50 drugs which are known to be used chronically (i.e., drugs which are commonly indicated for long-term use such as but not limited to antihypertensive, antilipidemic, and antidiabetic agents) rather than acutely (e.g., analgesics were excluded as they are mostly used acutely), and excluding

natural products and supplements, hormonal replacement treatments, eye drops, drugs in gel formulations, and inhalers. The list of the top 50 chronic drugs selected to be the target of our study are listed in Table 3 below.

Number	The selected drug
1	Simvastatin
2	Omeprazole
3	Bendroflumethiazide
4	Ramipril
5	Amlodipine
6	Atenolol
7	Atorvastatin
8	Lansoprazole
9	Metformin
10	Lisinopril
11	Amitriptyline
12	Ranitidine
13	Citalopram
14	Bisoprolol
15	Perindopril
16	Alendronate Sodium
17	Fluoxetine
18	Doxazosin
19	Allopurinol
20	Warfarin
21	Losartan
22	Gliclazide
23	Felodipine
24	Ezetimibe
25	Enalapril
26	Furosemide
27	Propranolol
28	Clopidogrel
29	Rosuvastatin
30	Quinine
31	Valsartan
32	Methotrexate
33	Esomeprazole
34	Pravastatin
35	Nifedipine
36	Gabapentin
37	Sertraline
38	Diltiazem
39	Venlafaxine
40	Isosorbide Mononitrate
41	Finasteride
42	Tamoxifen
43	Nicorandil
44	Diazepam
45	Mirtazapine
46	Sodium valproate
47	Digoxin
48	Carbamazepine
49	Pioglitazone
50	Pregabalin

Table 3: List of top 50 commonly used chronic drugs in the UK.

3.2 Selection of candidate genetic variants

Figure 15 below clarifies the selection process of candidate genetic variants used in this chapter and subsequent chapters.

3.2.1 Identifying the initial list of candidate SNPs.

In brief, we selected 34 genes from DrugBank that encoded all common drug-metabolizing enzymes or drug transporters (see chapter I) or were not very common but affect at least one of the 50 selected drugs (i.e. the drug is a substrate for the gene) ; 2 additional genes encoding relevant enzymes (PTGS1(COX-1) and CES1) were also included as they affect some important cardiovascular agents we selected in our study (see Table 4). 757 genetic variants in these 36 genes were extracted from PharmGKB *(https://www.pharmgkb.org/downloads).* The genotype data for 156 SNPs were not available in any of the cohorts: GoDARTs, GS, GoSHARE, and the UKBB as they don't exist in the European ancestry leaving us with a starting list of 601 SNPs.

				The selected genes					
ABCB1	*ABCC1*	*ABCC2*	*ABCC3*	*ABCC4*	*ABCC5*	*ABCG2*	*CYP1A1*	*CYP1A2*	*CYP1B1*
CYP2A6	*CYP2B6*	*CYP2C18*	*CYP2C19*	*CYP2C8*	*CYP2C9*	*CYP2D6*	*CYP2E1*	*CYP3A4*	*CYP3A5*
CYP3A7	*SLC22A1*	*SLC22A11*	*SLC22A2*	*SLC22A3*	*SLC22A4*	*SLC22A5*	*SLC22A6*	*SLC22A8*	*SLCO1A2*
SLCO1B1	*SLCO1B3*	*SLCO1C1*	*SLCO2B1*	*CES1*	*PTGS1*				

Table 4: List of 36 important selected genes which could affect drug response.

3.2.2 SNPs selection based on MAF (>=0.05) and HWE (p >= $1*10^{-8}$) criteria.

For the GoDARTs cohort, the genotype data were available for 509 out of 601 SNPs. The minor allele frequency (MAF) for each of these 509 SNPs was calculated with 320 SNPs fulfilling the MAF condition of >= 0.05 (retaining two SNPs with MAF below than but very close to 0.05 as these passed MAF condition in the other two cohorts). Two of these SNPs were not in HWE (rs113667357 and rs59502379, p<= $1*10^{-8}$) and were excluded.

For the GS cohort, the genotype data were available for 455 SNPs. 322 SNPs had a MAF >= 0.05 (retaining one SNP with MAF below than but very close to 0.05 as it passed MAF condition in the other two cohorts). All of the 322 SNPs passed HWE criteria.

For the GoSHARE cohort, the genotype data were available for 509 SNPs. 312 SNPs had a MAF >= 0.05. All of the 312 SNPs passed HWE criteria.

For the initial analyses, the three cohorts were combined together resulting in a 333 SNPs, although 2 SNPs (rs2070676 and rs2515641) were now out of HWE and were excluded resulting in a total of 331 SNPs.

3.2.3 SNPs selection based on LD test results ($r^2 < 0.5$).

The SNPs were then pruned to retain independent SNPs, with $r^2<0.5$. Initially, the pairwise correlation was assessed, with 53901 SNP pairs being considered independent, and 714 SNP pairs being correlated. To prune the 331 SNPs,

I first reviewed "Very Important Pharmacogenes (VIPs)" summaries from PharmGKP [128]. These summaries provide information on important variants which have been shown to affect drug response. PharmGKB included these variants based on many resources including the table of pharmacogenomic biomarkers issued by the Food and Drug Administration (FDA) (*https://www.fda.gov/drugs/science-and-research-drugs/table-pharmacogenomic-biomarkers-drug-labeling*), CPIC (*https://cpicpgx.org/alleles/*) , and the Pharmacogenomic Research Network (PGRN , *https://www.pgrn.org/*). After reviewing all VIPs summaries for all selected genes, I have identified 26 important genetic variants to be retained in my analysis (e.g., the known 'classic' variants such as *CYP2C9**2 and **3*). After comparing this list of 26 SNPs with the list of SNPs from the EU "Ubiquitous PGx (U-PGx)" (*https://upgx.eu/wp-content/uploads/2018/08/Deliverable-D2.1UPDATED_P6-KNMP.pdf*), I found that all of my selected SNPs are included in this list of SNPs (U-PGx) as long as they fulfil my selection criteria (MAF >= 5% in the European ancestry). In my list of 331 SNPs, I have identified 82 variants to be correlated to at least one of these 26 SNPs and were not recognised as important variants by PharmGKB. These have been removed from my analysis. Table 5 below illustrates more details on the 26 important variants selected in our analysis.

No.	Gene	rsid	* allele	Chromosome	Amino acid [codon] change *	Codon number (nucleotide change)*	Variant class *	Enzyme/transporter function *
1	*ABCB1*	rs1045642	*2/*13	7	I [ATT] > M [ATG]	3645(T>G)	Missense	Conflicting findings
2	*ABCB1*	rs1128503	NA	7	G [GGT] > G [GGC]	1446(T>C)	Synonymous	Conflicting findings
3	*ABCB1*	rs2032582	NA	7	S [TCT] > A [GCT]	2887(T>G)	Missense	Conflicting findings
4	*ABCG2*	rs2231142	NA	4	Q [CAG] > E [GAG]	421(C>G)	Missense	Decreased
5	*CYP1A2*	rs762551	*1F	15	NA	NA	Intron	Increased
6	*CYP2A6*	rs28399433	*9A/*9B/*13/*15	19	NA	NA	Upstream	Decreased
7	*CYP2A6*	rs8192726	*9b	19	NA	NA	Intron	Decreased
8	*CYP2B6*	rs2279343	*4	19	K [AAG] > R [AGG]	785(A>G)	Missense	Increased
9	*CYP2C19*	rs12248560	*17	10	NA	NA	Upstream	Increased
10	*CYP2C19*	rs4244285	*2	10	P [CCG] > P [CCA]	681(G>A)	Synonymous	No function
11	*CYP2C8*	rs10509681	*3	10	K [AAA] > R [AGA]	1196(A>G)	Missense	Increased
12	*CYP2C8*	rs1058930	*4	10	I [ATC] > I [ATT]	792(C>T)	Missense	Decreased
13	*CYP2C8*	rs11572080	*3	10	R [AGG] > M [ATG]	416(G>T)	Missense	Increased
14	*CYP2C8*	rs17110453	*1C	10	NA	NA	Upstream	Conflicting findings
15	*CYP2C8*	rs7909236	*1B	10	NA	NA	Upstream	Increased
16	*CYP2C9*	rs1057910	*3	10	I [ATT] > L [CTT]	1075(A>C)	Missense	No function
17	*CYP2C9*	rs1799853	*2	10	R [CGT] > C [TGT]	430(C>T)	Missense	Decreased
18	*CYP2D6*	rs1065852	*10	22	P [CCA] > S [TCA]	100(C>T)	Missense	Decreased
19	*CYP2D6*	rs1135840	*2	22	S [AGC] > T [ACC]	1457(G>C)	Missense	Normal function
20	*CYP2D6*	rs16947	*2	22	R [CGC] > C [TGC]	886(C>T)	Missense	Normal function
21	*CYP2D6*	rs28371725	*41	22	NA	NA	Intron	Decreased
22	*CYP2D6*	rs3892097	*4	22	NA	NA	Splice Acceptor	No function
23	*CYP3A4*	rs2242480	*1G	7	NA	NA	Intron	Increased
24	*SLC22A1*	rs12208357	NA	6	R [CGC] > C [TGC]	181(C>T)	Missense	Decreased
25	*SLCO1B1*	rs2306283	*1B	12	N [AAT] > H [CAT]	388(A>C)	Missense	Increased
26	*SLCO1B1*	rs4149056	*5	12	V [GTG] > A [GCG]	521(T>C)	Missense	Decreased

Table 5: List of 26 important variants involved in drug response as defined by their molecular characteristics.

* *According to https://www.ncbi.nlm.nih.gov/snp/docs/refsnp_report/helpdoc/.*

* *Amino acids: **I** = Isoleucine , **M** = Methionine , **G** = Glycine , **S** = Serine , **A** = Alanine , **Q** = Glutamine , **E** = Glutamate , **K** = Lysine , **R** = Arginine , **P** = Proline , **L** = Leucine , **C** = Cysteine , **T** = Threonine , **N** = Asparagine , **H** = Histidine , **V** = Valine .*

* *Each codon consists of [3 nucleotides] encoding for a specific amino acid. There are 4 nucleotides: **A** = Adenine , **T** = Thymine , **G** = Guanine , **C** = Cytosine.*

* *According to both https://www.pharmvar.org/genes and https://www.pharmgkb.org/vips .*

Secondly, I reviewed the remaining 223 SNPs for pairwise correlation and retained 76 independent SNPs. Finally, for the correlated SNP pairs (where there was no reason to choose one over the other) one SNP was selected at random among any group of correlated SNPs (i.e. the SNP might be correlated with one or more SNPs) resulting in 60 SNPs retained as proxies. Overall, this resulted in 162 independent SNPs (26 + 76 + 60) were used in the analysis.

Having identified our list of candidate variants for use in the combined cohort, I then extracted the genotype data for these 162 SNPs for only white British individuals in the UKBB cohort (n=408091). Of note, 9 SNPs were found to be of HWE including the important variants rs28399433 (*CYP2A6**9A), rs16947(*CYP2D6**2), rs3892097 (*CYP2D6**4), rs1065852 (*CYP2D6**10), and rs1135840 (*CYP2D6*2*).

However, these 9 SNPs were included in the UKBB analysis for the purpose of keeping the consistency, so we have the same list of SNPs across all cohorts.

The final selected 162 SNPs with their MAF and HWE results from all cohorts, as classified into the 3 categories of SNPs presented in figure 15 below, are shown in *supplementary material 2.*

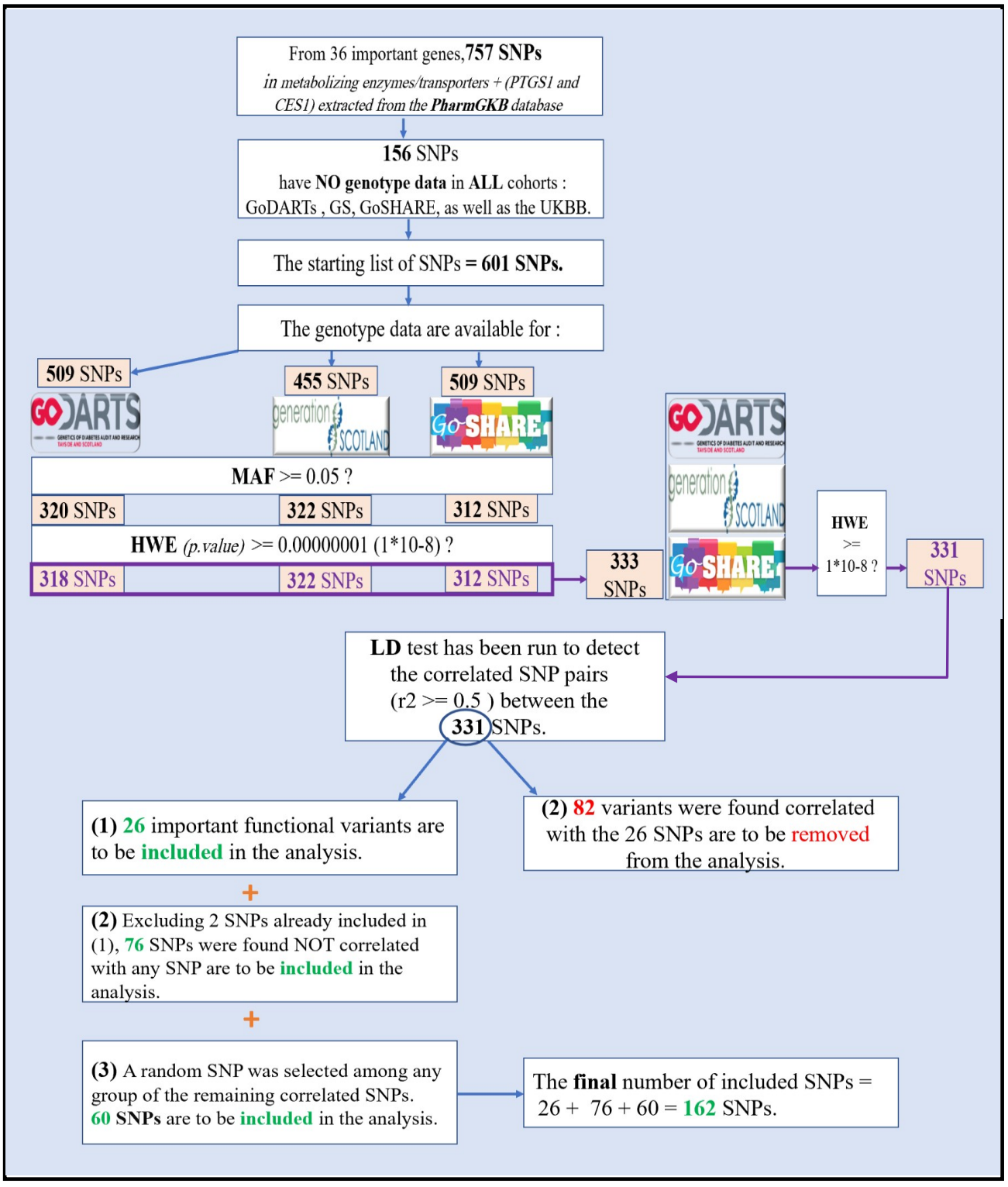


Figure 15: A summary workflow for the process of selection candidate genetic variants for the combined cohort.

3.3 Defining drug response phenotypes

Drug response phenotypes were defined depending on the nature of the prescribing data available to us at the time of this study. The approach is outlined in detail in chapter II, and therefore only briefly reviewed here.

In the cross-sectional UKBB data, we looked for a deviation of genotype frequencies in those prescribed a particular drug from the total UKBB population. In this way, two phenotypes were defined.

- *Increased drug tolerability phenotype*: this occurs when the variant allele found to be highly distributed within the drug users compared to non-users. If those with a specific drug-variant pair are higher, this will reflect that this pair was advantageous for these patients in term of increased efficacy, low side effects, or both.
- *Decreased drug tolerability phenotype*: this occurs when the variant allele found to be significantly lower within the drug users compared to non-users. If those with a specific drug-variant pair are lower; this would reflect that this pair was disadvantageous for these patients in term of decreased efficacy, increased side effects, or both.

Importantly, deviation of genotype frequencies can occur due to the underlying disease for which a drug is prescribed and therefore not reflect drug tolerability. Therefore, for a variant to be considered potentially associated with drug tolerability, it had to not be associated with any known disease traits, which we assessed by undertaking a phenome-wide association study (PheWAS) in gene GeneAtlas (chapter II).

In the Scottish longitudinal data, we defined the generic "drug-stop" and "dose-decrease" phenotypes (as described in chapter II). Stopping the drug would indicate experiencing intolerable side effects, lack of therapeutic efficacy, or both. Decreasing the dose may be a sign of toxicity or extreme increase in the efficacy, which would be harmful such as sulphonylurea-induced hypoglycemia or warfarin-induced bleeding. These phenotypes are more likely to occur among poor CYP2C9 metabolizers. In chapter I, for example, I have highlighted that *CYP2C9*2/*3* carriers required lower warfarin doses to overcome the bleeding risk of this anticoagulant treatment.

3.4 Testing the association between the genetic variants and the phenotypes

Users of each of the selected 50 chronic drugs from both the UKBB cross-sectional prescribing data and the combined cohort longitudinal prescribing data (including patients from all of the 3 cohorts (i.e. GoDARTs + GS + GoSHARE)) were extracted. The case groups from the UKBB were users of a certain drug while controls are those not on the drug under investigation. Regarding the combined cohort, for each of the 50 drugs, users of the drug of interest are divided into case and control groups. The case groups are those who stop the drug or those who reduce their daily dose. The control groups are those who have never experienced any of these two phenotypes, respectively.

A case-control genetic analysis using a log-additive genetic model was run to explore the associations between our 162 selected genetic variants and each phenotype under each cohort for all 50 drugs. Given that each drug phenotype was evaluated for 162 independent SNPs, a Bonferroni adjusted p-value for significance was set at <= 0.00030 (0.05/162).

3.5 Development of an online database to view the results

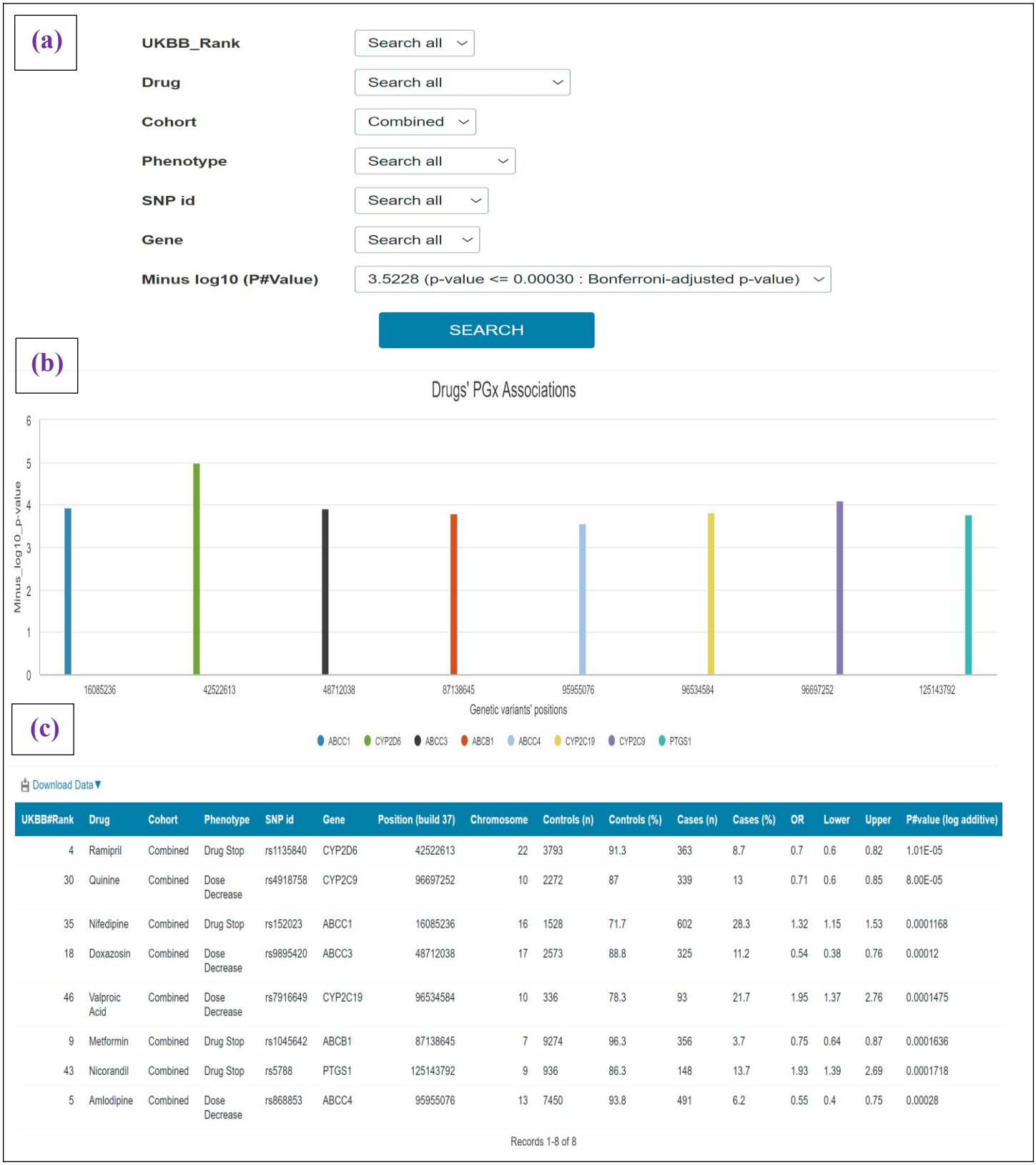
Given the large number of results produced, a dynamic, user-friendly online database was developed to view all results visually (graphs) or as tables using Caspio software (https://www.caspio.com/). The user can view and download the results according to the criteria of interest.

4. Results

In the combined Scottish data, examining the association between the drug-stop phenotype for 50 drugs and 162 genetic variants produced a total of 8,100 results; the same number of findings were also generated for the dose-decrease phenotype from the combined cohort taking the total findings into 16,200. For the UKBB cross-sectional cohort, again, 162 variants were compared for 50 drugs producing a total of 8,100 findings.

In order to facilitate viewing the results, all results can be accessed via an online database under this link:
https://c1abo933.caspio.com/dp/d81f7000c3c1854d29104a49b1d8

This application enables the user to select the results of interest from a search form. The results can be selected by drug rank, drug name, cohort, phenotype, SNP rsid, gene, and/or the p-value threshold. The results are shown in two formats: a dynamic summary diagram showing the top hits (genetic variants) according to the p-values and a detailed report table for the full results. The user can also download the searched results. Figure 16 below shows some screenshots from the application.



UKBB#Rank	Drug	Cohort	Phenotype	SNP id	Gene	Position (build 37)	Chromosome	Controls (n)	Controls (%)	Cases (n)	Cases (%)	OR	Lower	Upper	P#value (log additive)
4	Ramipril	Combined	Drug Stop	rs1135840	CYP2D6	42522613	22	3793	91.3	363	8.7	0.7	0.6	0.82	1.01E-05
30	Quinine	Combined	Dose Decrease	rs4918758	CYP2C9	96697252	10	2272	87	339	13	0.71	0.6	0.85	8.00E-05
35	Nifedipine	Combined	Drug Stop	rs152023	ABCC1	16085236	16	1528	71.7	602	28.3	1.32	1.15	1.53	0.0001168
18	Doxazosin	Combined	Dose Decrease	rs9895420	ABCC3	48712038	17	2573	88.8	325	11.2	0.54	0.38	0.76	0.00012
46	Valproic Acid	Combined	Dose Decrease	rs7916649	CYP2C19	96534584	10	336	78.3	93	21.7	1.95	1.37	2.76	0.0001475
9	Metformin	Combined	Drug Stop	rs1045642	ABCB1	87138645	7	9274	96.3	356	3.7	0.75	0.64	0.87	0.0001636
43	Nicorandil	Combined	Drug Stop	rs5788	PTGS1	125143792	9	936	86.3	148	13.7	1.93	1.39	2.69	0.0001718
5	Amlodipine	Combined	Dose Decrease	rs868853	ABCC4	95955076	13	7450	93.8	491	6.2	0.55	0.4	0.75	0.00028

Records 1-8 of 8

Figure 16: Screenshots from our online application showing 3 main parts for our drug-gene interactions results. For example, searching the results using specific criteria such as results significant after Bonferroni correction from the combined cohort **(a)** shows 8 hits **(b)** with the detailed report of results **(c)**.

I have identified 815 drug-gene associations with a nominal significance level ($p <= 0.05$) from the combined cohort considering both drug-stop and dose-decrease phenotypes. In the UKBB cross-section cohort, 544 findings were identified at this nominal p-value.

For a better representation of the results, we classified the results into 3 categories: 1) the associations which have passed Bonferroni significance level; 2) the known drug-gene associations which have been previously reported; and 3) potentially important associations which didn't pass Bonferroni significance level.

I present results for each category below.

4.1 Significant associations after Bonferroni correction ($p <=0.0003$)

Table 6, at the end of the results section, summarizes all associations under this category.

4.1.1 The combined cohort results (n= 8).

8 novel drug-gene associations passed the Bonferroni significance level. Below I begin with a brief description of these findings before showing the replication results from the UKBB primary care data later in chapter V.

4.1.1.1 Ramipril-rs1135840 (CYP2D6)

The rs1135840 (G >C, *CYP2C6*2*) variant at the *CYP2D6* gene was associated with being less likely to stop ramipril treatment ($p=1.01 \times 10^{-5}$). The C allele is associated with 30% (95%CI 18-40%) reduced risk to stop the treatment.

4.1.1.2 Quinine-rs4918758 (CYP2C9)

Carriers of the C allele at rs4918758 (T>C) variant in CYP2C9 had a 29% (95% CI 15-40%) lower tendency to reduce daily quinine dose per allele ($p=8 \times 10^{-5}$).

4.1.1.3 Nifedipine-rs152023 (ABCC1)

Carriers of the minor allele (C) at rs152023 in ABCC1 were 1.32 (1.15 – 1.53) times more likely to stop nifedipine treatment per allele compared to TT individuals ($p=1.168 \times 10^{-4}$).

4.1.1.4 Doxazosin-rs9895420 (ABCC3)

The odds of decreasing the daily dose of doxazosin in those carrying the minor (A) allele in rs9895420 *ABCC3* SNP was reduced by 46% (95% CI 24-62%) per allele compared to non-carriers (TT individuals) ($p=1.2 \times 10^{-4}$).

4.1.1.5 Valproic acid- rs7916649 (CYP2C19)

The A allele at rs7916649 (G>A) in CYPC19 was associated with 1.95 (1.37-2.76) times increased risk for daily dose reduction of valproic acid ($p=1.475 \times 10^{-4}$).

4.1.1.6 Metformin-rs1045642 (ABCB1)

The rs1045642 (A>G) mutation in ABCB1 transporter was associated with a 25% (95% CI 13-36%) lower odds of stopping metformin treatment per allele compared to wild-type individuals ($p=1.636 \times 10^{-4}$).

4.1.1.7 Nicorandil-rs5788 (PTGS1)

Those with an A-allele at rs5788 (C>A) variant in *PTGS1* gene were 1.93 times (95% CI 1.39-2.69) more likely to stop nicorandil treatment per allele ($p=1.718 \times 10^{-4}$).

4.1.1.8 Amlodipine-rs868853 (ABCC4)

There was a large impact of the minor allele (C) at rs868853 variant on amlodipine daily dose. Individuals carrying the C allele were 45 % (95%CI 25-60%) less likely to decrease amlodipine daily dose per allele as compared to non-carriers (TT individuals) ($p=2.8 \times 10^{-4}$).

4.1.2 The UKBB (cross-sectional) cohort results (n=4)

25 drug-gene pairs reached the Bonferroni significance threshold. However, as discussed, these may well represent disease associations rather than drug tolerance. Using GeneAtlas we identified 21 of these 25 variants were significantly associated with the disease for which the drug was indicated, and these are not discussed. 4 drug-variant associations were not disease associations and may therefore represent drug tolerance phenotypes. I describe the results of these four findings below.

4.1.2.1 Lansoprazole-rs555754 (SLC22A3)

The A allele at rs555754 (G>A) variant in *SLC22A3* gene was significantly lower amongst lansoprazole users compared to non-users ($p=1.58 \times 10^{-5}$). It is 5% (3%-7%) less likely for the A allele to be seen amongst lansoprazole users compared to non-users.

4.1.2.2 Bendroflumethiazide-rs3743527 (ABCC1)

The T allele at rs3743527 (C>T) of ABCC1 was lower in those treated with bendroflumethiazide ($p=4.87 \times 10^{-5}$) with this allele being 5% (2%-7%) less likely to be detected in this group than the rest of UK Biobank.

4.1.2.3 Gabapentin-rs8187843 (ABCC1)

The ABCC1 transporter also appeared to be associated with gabapentin use. The A allele at rs8187843 (G>A) was at a lower frequency in the gabapentin-treated group ($p=2.12 \times 10^{-4}$) compared to the rest of UK Biobank. There was a 22% (11%-32%) lower odds of observing this variant in this treatment group compared to the non-treatment group.

4.1.2.4 Rosuvastatin-rs2231135 (ABCG2)

The rs2231135 (A>G) variant in the *ABCG2* gene was associated with rosuvastatin use (p=3.0 × 10-4). The G allele was more frequent amongst individuals on this drug, with a 22% (1%-35%) higher chance to see this variant within rosuvastatin users compared to non-users.

4.2 The associations which have been previously reported

The number of nominally significant results (p <= 0.05) totalled 815 from the combined cohort and 544 from the UKBB cohort. The PharmGKB database was reviewed for all of these findings to identify drug-variant pairs which have been previously reported. A literature search using PubMed was also undertaken to identify similar studies. I found that the majority of our associations have not been reported before (and are likely to be false-positive findings at this significance level). However,

there were 24 associations (15 from the combined cohort and 9 from the UKBB cohort) where at least one paper has been published previously studying the same drug-variant pair. 19 of our results (11 from the combined and 8 from the UKBB cohorts) appear to be consistent with the previous findings increasing their prior likelihood of a true association. 4 findings from the combined cohort and 1 other result from the UKBB cohort were not consistent with our results. Here I focus on the results that were consistent with prior literature. As we are looking for replicated but not novel associations, a p-value of <= 0.05 was considered sufficient; however, this can only be considered exploratory, and further replication would be required of these results.

Table 6, at the end of the results section, summarizes all associations under this category.

4.2.1 The combined Scottish cohort results (n = 11).

4.2.1.1 Gliclazide-rs1057910 (CYP2C9)

The rs1057910 (A>C, *CYP2C9*3*) polymorphism has been reported to be associated with increased risk of sulfonylurea-induced hypoglycemia (OR = 1.68, p=0.011) [129]. Consistent with this, the carriers of this variant were 26% (7%-48%) more likely to decrease their gliclazide daily dose per allele (p=0.00696).

4.2.1.2 Simvastatin-rs2231142 (ABCG2)

The rs2231142 (G>T) variant in the *ABCG2* gene has been associated with a 46% reduction in simvastatin clearance (p=0.017) [130]. Our results show that this variant is associated with 18% (p=0.0069) lower odds of stopping simvastatin treatment.

4.2.1.3 Methotrexate-rs9895420 (ABCC3)/rs1128503 (ABCB1)

The rs9895420 (T>A) *ABCC3* SNP has been reported to be correlated with reduced methotrexate efficacy and side effects (p=0.01 and 0.06, respectively) [131]. Our findings show that this variant is correlated with being 90% (95%CI 25-99%) less likely to stop methotrexate treatment (p=0.0134). In addition, the minor allele (A) at the rs1128503 (A>G) SNP in the ABCB1 transporter was previously reported to be associated with increased methotrexate toxicity [132]. Consistent with this, our results show that carriers of the A allele were 1.68 (1.07-2.63) times more likely to stop their methotrexate treatment per allele compared to non-carriers (p=0.0227).

4.2.1.4 Carbamazepine-rs762551(CYP1A2)/ rs2242480(CYP3A4)/ rs1128503(ABCB1) (3 associations)

It has been previously reported that carriers of the C allele at rs762551 (C(*minor allele*)>A) variant in the *CYP1A2* gene experienced lower carbamazepine clearance and increased concentrations (p =0.004) [133]. In the combined Scottish cohorts, I also show that this allele is associated with a 1.33 (1.05-1.67) times greater tendency to decrease carbamazepine daily dose (p=0.0168).

It has also been reported that the T allele at rs2242480 (C>T) variant in the *CYP3A4* gene is associated with decreased carbamazepine concentrations (p =0.027) [134]. Our results show that this allele is also linked with a 37% lower odds to decrease the daily dose (p=0.0368).

Finally, the A allele at rs1128503 (A(*minor allele*)>G) variant in *ABCB1* gene was associated with increased carbamazepine clearance in African American but not Caucasian population (p= 0.036) [135]. However, I observed a weak association

between this variant and being 18% (1%-38%) more likely to stop the drug (p=0.047) in a British cohort.

4.2.1.5 Pioglitazone-rs10509681 (CYP2C8)

The rs10509681 (T>C, *CYP2C8*3*) variant has been reported to be associated with a 29.7% lower pioglitazone plasma levels (p=0.01) [136]. Consistently, our results show that this variant allele is correlated with 33% (3%-70%) increased risk of stopping the drug (p=0.026).

4.2.1.6 Warfarin-rs2242480 (CYP3A4)

In a previous study [137], the rs2242480 (C>T) *CYP3A4* SNP has been associated with increased warfarin clearance. In our study, I found that this SNP is also correlated with 58% (6%-134%) increased odds to stop warfarin therapy (p=0.032).

4.2.1.7 Atorvastatin-rs2032582 (ABCB1)

The A allele at rs2032582 (A(*minor allele*)>T) *ABCB1* variant has been shown to be connected with increased atorvastatin efficacy [138]. Consistently, our results reveal that there is 15% (1%-53%) lower likelihood for the A allele carriers to stop atorvastatin treatment per allele (p=0.035).

4.2.1.8 Clopidogrel-rs1057910 (CYP2C9)

It has been reported that the *CYP2C9* mutation, rs1057910 (A>C, *CYP2C9*3*), is linked with decreased clopidogrel efficacy (p=0.045) [139]. I have also noted that this variant allele (C allele) to be correlated with 86% (21%-98%) lower odds of decreasing daily clopidogrel dose (p=0.036).

4.2.2 The UKBB (cross-sectional) cohort results (n=8).

4.2.2.1 Amitriptyline-rs1065852 (CYP2D6)

The rs1065852 (G>A, *CYP2D6*10*) SNP has been consistently correlated with poor amitriptyline clearance and increased toxicity [140]. Our study shows that the distribution of the A allele is significantly lower in the amitriptyline treated group compared to non-treatment group (p=0.00075) with a 6% (3%-10%) lower chance to observe this allele among amitriptyline treated individuals.

4.2.2.2 Amlodipine-rs1045642 (ABCB1)

There was a slight reduction in G allele frequency at rs1045642 (A>G) in ABCB1 has in amlodipine users (OR = 0.97(0.95-0.99), p=0.00452). This is supported by a previous study in which this variant was associated with decreased amlodipine clearance and increased concentrations [141].

4.2.2.3 Simvastatin-rs4149056 (SLCO1B1)

The rs4149056 (T>C, *SLCO1B1*5*) is a common SNP which has been frequently associated with simvastatin toxicity [142]. The C-allele was underrepresented in the simvastatin users with an OR of 0.98 (p=0.016) in the UKBB cohort treated with simvastatin.

4.2.2.4 Clopidogrel-rs12248560/rs4244285 (CYP2C19) (2 associations)

Two common *CYP2C19* SNPs, rs12248560 (C>T, *CYP2C19*17*) and rs4244285 (G>A, *CYP2C19*2*), have been reported to be correlated with increased [143] and decreased [144] clopidogrel efficacy respectively. In the UKBB cohort, there was a significant increase (OR = 1.08(1.01-1.14), p=0.024) and reduction

(OR =0.92(0.86-1), p=0.039) in the distribution of the T and A alleles respectively among clopidogrel users.

4.2.2.5 Citalopram-rs28371725 (CYP2D6)

The rs28371725 (C>T, *CYP2D6*41*) variant has been linked to higher citalopram efficacy [145]. There was a slight increase in T allele frequency amongst patients on citalopram when compared to non-users (OR = 1.06 (1.01-1.13), p=0.026).

4.2.2.6 Pioglitazone-rs10509681 (CYP2C8)

As described above, there was an association between the rs10509681 (T>C, *CYP2C8*3*) variant and reduced pioglitazone plasma levels and increased the chance of stopping the drug in the combined Scottish cohorts. Interestingly, a consistent association was also seen in the UKBB cohort as the C-allele was underrepresented among pioglitazone users (OR = 0.9 (0.82-1), p=0.043).

4.2.2.7 Fluoxetine-rs1065852 (CYP2D6)

The rs1065852 (G>A, *CYP2D6*10*) variant has been associated with increased fluoxetine AUC [146]. This SNP was also noted with lower distribution within fluoxetine users from the UKBB (OR = 0.96 (0.91-1), p=0.048).

4.3 Potential important associations which didn't pass Bonferroni significance level.

Under this category of results, I include associations, which although not significant after Bonferroni correction, are still of interest as they fulfil 3 criteria together: Firstly, the p-value significance level is moderate (> 0.00030 and <= 0.009). Secondly, they occur in an enzyme/transporter known as a substrate for the drug (according to DrugBank database (described in chapter II) or external references). Thirdly, they also occur with a genetic variant has been previously associated with at least one drug response phenotype (according to PharmGKB). I have recognized 12 findings belong to this category from the combined cohort and other 8 results from the UKBB cohort, which are summarized in Table 6 at the end of the results section. In the below sections, I describe these findings.

4.3.1 The combined cohort results (n = 12).

4.3.1.1 Clopidogrel-rs12353214 (PTGS1)

In an interesting association with a large effect size, the T allele at rs12353214 (C>T) *PTGS1 (COX-1)* variant was associated with a 43% (95%CI 20%-59%) reduced risk to stop the treatment ($p=5.3 \times 10^{-4}$).

4.3.1.2 Valsartan-rs4918758 (CYP2C9)

Carriage of the rs4918758 (T>C) *CYP2C9* variant was associated with 33% (15%-47%) lower tendency to decrease daily valsartan dose ($p=5.5 \times 10^{-4}$).

4.3.1.3 Ezetimibe-rs3842 (ABCB1)

Patients harbouring the rs3842 (T>C) polymorphism in the *ABCB1* gene were seen 89% (34%-167%) more likely to stop ezetimibe treatment per allele ($p=5.6 \times 10^{-4}$).

4.3.1.4 Enalapril-rs2244614 (CES1)

Carrying the G allele at rs2244614 (G(*minor allele*)>A) mutation in *CES1* gene is associated with being 38% (17%-53%) at lower chance to decrease daily enalapril dose ($p=8.73 \times 10^{-4}$).

4.3.1.5 Lansoprazole- rs9282564 (ABCB1)

Lansoprazole users carrying the C allele at rs9282564 (T>C) *ABCB1* SNP had 18% (7%-30%) lower likelihood to reduce their daily dose per allele (p=0.0027).

4.3.1.6 Atenolol-rs628031 (SLC22A1)

Harbouring the A allele at rs628031(A(*minor allele*)>G) *SLC22A1* variant was associated with 21% (7%-38%) increased risk to stop atenolol treatment (p=0.0034).

4.3.1.7 Digoxin-rs4728709 (ABCB1)

Our results also show that the rs4728709 (G>A) *ABCB1* variant is correlated with 46% (15%-66%) decreased chance to reduce daily digoxin dose (p=0.0045).

4.3.1.8 Nifedipine-rs4728709 (ABCB1)

The same variant (rs4728709) was associated with a 40% (16%-57%) lower odds to stop nifedipine treatment (p=0.0063).

4.3.1.9 Ranitidine-rs316019 (SLC22A2)

Carriers of the A allele at rs316019 (A(*minor allele*)>C) *SLC22A2* SNP were at 15% (4%- 25%) lower risk to reduce the daily dose of ranitidine per allele (p=0.0065).

4.3.1.10 Valproic acid-rs2279343 (CYP2B6)

Individuals with rs2279343 (A>G) *CYP2B6* polymorphism are at 69% (15%-148%) increased tendency to drop their valproic acid daily dose per allele (p=0.0082).

4.3.1.11 Clopidogrel-rs3815583 (CES1)

A *CES1* SNP, rs3815583 (A>C), was associated with 65% (13%-86%) lower chance to decrease the daily dose of clopidogrel (p=0.0083).

4.3.1.12 Simvastatin-rs1080985(CYP2D6)

The *CYP2D6* SNP, rs1080985 (G>C), has shown 16% (4%-27%) reduced odds of decreasing the daily simvastatin dose (p=0.0089).

4.3.2 The UKBB (cross-sectional) cohort results (n=8).

4.3.2.1 Atorvastatin-rs17287570 (ABCC1)

The C allele at rs17287570 (A>C) variant in ABCC1 transporter has 5% (2%-8%) lower odds to be detected within the atorvastatin treated group compared to the non-treatment group (p=0.000918).

4.3.2.2 Esomeprazole-rs12248560(CYP2C19)/rs1128503/rs1045642(ABCB1) (3 associations)

While the minor alleles at rs12248560 (C>T) *CYP2C19* and rs1128503 (A(*minor allele*)>G) *ABCB1* SNPs had 12% (4%-21% , p=0.0023)) and 9% (2%-16% , p=0.0063) higher chance to be recognized among patients treated with esomeprazole

respectively, the G allele at rs1045642 *ABCB1* (A>G) polymorphism has shown 8% (2%-14%) lower tendency to be identified in the esomeprazole-treated group (p=0.0086).

4.3.2.3 Methotrexate-rs4793665 (ABCC3)

In an association with very small effect size, the C allele at rs4793665 (C(*minor allele*)>G) *ABCC3* variant has only 1% (3%-16%) less likelihood to be observed among patients on methotrexate therapy (p=0.0026).

4.3.2.4 Allopurinol-rs2231135 (ABCG2)

The distribution of the G allele at rs2231135 (A>G) variant in ABCG2 transporter was significantly lower among individuals on allopurinol (p=0.0035). We have 13% (4%-21%) lower odds to see this variant in this treatment group.

4.3.2.5 Enalapril-rs7317112 (ABCC4)

The ABCC4 transporter SNP, rs7317112 (A>G), is significantly highly distributed within patients on enalapril (p=0.00455) with the odds being 8% (3%-14%) higher in the treatment group.

4.3.2.6 Clopidogrel-rs9332197 (CYP2C9)

Carriers of the rs9332197 (T>C) *CYP2C9* variant were significantly lower in the clopidogrel-treatment group compared to the non-treatment group (p=0.0062). There is 15% (4%-25%) decreased likelihood for this SNP to be detected in the treatment group.

Table 6 below summarizes our important drug-gene interaction findings as classified into the 3 categories I mentioned above.

			1) Results significant after Bonferroni correction						
No.	**Drug**	**Cohort**	**Phenotype**	**SNP_id**	**Gene**	**OR**	**Lower**	**Upper**	**P.value**
1	Ramipril	Combined	Drug Stop	rs1135840 (*2D6*2*)	*CYP2D6*	0.7	0.6	0.82	1.01E-05
2	Quinine	Combined	Dose Decrease	rs4918758	*CYP2C9*	0.71	0.6	0.85	8.00E-05
3	Nifedipine	Combined	Drug Stop	rs152023	*ABCC1*	1.32	1.15	1.53	0.000117
4	Doxazosin	Combined	Dose Decrease	rs9895420	*ABCC3*	0.54	0.38	0.76	0.00012
5	Valproic Acid	Combined	Dose Decrease	rs7916649	*CYP2C19*	1.95	1.37	2.76	0.000148
6	Metformin	Combined	Drug Stop	rs1045642	*ABCB1*	0.75	0.64	0.87	0.000164
7	Nicorandil	Combined	Drug Stop	rs5788	*PTGS1*	1.93	1.39	2.69	0.000172
8	Amlodipine	Combined	Dose Decrease	rs868853	*ABCC4*	0.55	0.4	0.75	0.00028
9	Lansoprazole	UKBB	NA	rs555754	*SLC22A3*	0.95	0.93	0.97	1.58E-05
10	Bendroflumethiazide	UKBB	NA	rs3743527	*ABCC1*	0.95	0.93	0.98	4.87E-05
11	Gabapentin	UKBB	NA	rs8187843	*ABCC1*	0.78	0.68	0.89	0.000212
12	Rosuvastatin	UKBB	NA	rs2231135	*ABCG2*	1.22	1.1	1.35	0.000302
			2) Results replicating previous studies' findings						
No.	**Drug**	**Cohort**	**Phenotype**	**SNP_id**	**Gene**	**OR**	**Lower**	**Upper**	**P.value**
1	Gliclazide	Combined	Dose Decrease	rs1057910 (*2C9*3*)	*CYP2C9*	1.26	1.07	1.48	0.00696
2	Simvastatin	Combined	Drug Stop	rs2231142	*ABCG2*	0.82	0.7	0.95	0.007
3	Methotrexate	Combined	Drug Stop	rs9895420	*ABCC3*	0.1	0.01	0.74	0.01347
4	Methotrexate	Combined	Drug Stop	rs1128503	*ABCB1*	1.68	1.07	2.63	0.0227
5	Carbamazepine	Combined	Dose Decrease	rs762551 (C(*minor*)>A)	*CYP1A2*	1.33	1.05	1.67	0.01682
6	Carbamazepine	Combined	Dose Decrease	rs2242480	*CYP3A4*	0.63	0.39	1	0.03689
7	Carbamazepine	Combined	Drug Stop	rs1128503	*ABCB1*	1.18	1	1.38	0.04734
8	Pioglitazone	Combined	Drug Stop	rs10509681 (*2C8*3*)	*CYP2C8*	1.33	1.03	1.7	0.02693
9	Warfarin	Combined	Drug Stop	rs2242480	*CYP3A4*	1.58	1.06	2.34	0.03297
10	Atorvastatin	Combined	Drug Stop	rs2032582	*ABCB1*	0.85	0.74	0.99	0.03545
11	Clopidogrel	Combined	Dose Decrease	rs1057910 (*2C9*3*)	*CYP2C9*	0.14	0.02	0.97	0.03625
12	Amitriptyline	UKBB	NA	rs1065852 (*2D6*10*)	*CYP2D6*	0.94	0.9	0.97	0.00075
13	Amlodipine	UKBB	NA	rs1045642	*ABCB1*	0.97	0.95	0.99	0.00452
14	Simvastatin	UKBB	NA	rs4149056 (*1B1*5*)	*SLCO1B1*	0.98	0.96	1	0.01621
15	Clopidogrel	UKBB	NA	rs12248560 (*2C19*17*)	*CYP2C19*	1.08	1.01	1.14	0.02417
16	Clopidogrel	UKBB	NA	rs4244285 (*2C19*2*)	*CYP2C19*	0.92	0.86	1	0.039
17	Citalopram	UKBB	NA	rs28371725 (*2D6*41*)	*CYP2D6*	1.06	1.01	1.13	0.02689
18	Pioglitazone	UKBB	NA	rs10509681 (*2C8*3*)	*CYP2C8*	0.9	0.82	1	0.0436
19	Fluoxetine	UKBB	NA	rs1065852 (*2D6*10*)	*CYP2D6*	0.96	0.91	1	0.04845
			3) Potential important associations						
No.	**Drug**	**Cohort**	**Phenotype**	**SNP_id**	**Gene**	**OR**	**Lower**	**Upper**	**P.value**
1	Clopidogrel	Combined	Drug Stop	rs12353214	*PTGS1*	0.57	0.41	0.8	0.00053
2	Valsartan	Combined	Dose Decrease	rs4918758	*CYP2C9*	0.67	0.53	0.85	0.00055
3	Ezetimibe	Combined	Drug Stop	rs3842	*ABCB1*	1.89	1.34	2.67	0.000569
4	Enalapril	Combined	Dose Decrease	rs2244614	*CES1*	0.62	0.47	0.83	0.000874
5	Lansoprazole	Combined	Dose Decrease	rs9282564	*ABCB1*	0.81	0.7	0.93	0.00275
6	Atenolol	Combined	Drug Stop	rs628031	*SLC22A1*	1.21	1.07	1.38	0.00345
7	Digoxin	Combined	Dose Decrease	rs4728709	*ABCB1*	0.54	0.34	0.85	0.00458
8	Nifedipine	Combined	Drug Stop	rs4728709	*ABCB1*	0.6	0.43	0.84	0.00633
9	Ranitidine	Combined	Dose Decrease	rs316019	*SLC22A2*	0.85	0.75	0.96	0.00654
10	Valproic Acid	Combined	Dose Decrease	rs2279343 (2B6*4)	*CYP2B6*	1.69	1.15	2.48	0.00821
11	Clopidogrel	Combined	Dose Decrease	rs3815583	*CES1*	0.35	0.14	0.87	0.00833
12	Simvastatin	Combined	Dose Decrease	rs1080985	*CYP2D6*	0.84	0.73	0.96	0.00895
13	Atorvastatin	UKBB	NA	rs17287570	*ABCC1*	0.95	0.92	0.98	0.000919
14	Esomeprazole	UKBB	NA	rs12248560	*CYP2C19*	1.12	1.04	1.21	0.002366
15	Esomeprazole	UKBB	NA	rs1128503	*ABCB1*	1.09	1.02	1.16	0.006394
16	Esomeprazole	UKBB	NA	rs1045642	*ABCB1*	0.92	0.86	0.98	0.008608
17	Methotrexate	UKBB	NA	rs4793665	*ABCC3*	1.1	1.03	1.16	0.002604
18	Allopurinol	UKBB	NA	rs2231135	*ABCG2*	0.87	0.79	0.96	0.00357
19	Enalapril	UKBB	NA	rs7317112	*ABCC4*	1.08	1.03	1.14	0.004557
20	Clopidogrel	UKBB	NA	rs9332197	*CYP2C9*	0.85	0.75	0.96	0.006261

Table 6: Summary of 51 most important drug-gene associations as classified into 3 categories.

5. Discussion

To the best of our knowledge, this is the first pharmacogenomic study covering a large variety of commonly used chronic drugs (n = 50) in the UK and study the influence of 162 different genetic variants in important enzymes and transporters on 3 drug response phenotypes (drug-stop, dose-decrease, and genotype distribution changes) for each drug. A total of 16,200 (combined cohort) and 8,100 (UKBB cohort) results have been generated and are available to view from an online database. Below I discuss our associations as classified into three categories.

5.1 Significant associations after Bonferroni correction.

5.1.1 The combined cohort results (n = 8)

The 8 significant associations under this section will be investigated further in Chapter V. In Chapter V; I will use the UKBB primary care data, which became available towards the end of my PhD, to check for the replicability of these 8 associations and provide a detailed discussion for each finding.

5.1.2 The UKBB (cross-sectional) cohort results (n = 4)

5.1.2.1 Lansoprazole-rs555754 (SLC22A3)

The rs555754 (G>A) variant in SLC22A3 (OCT3) transporter was modestly underrepresented among lansoprazole users with only 5% difference as compared to non-users. This transporter is expressed in the hepatic basolateral membrane facilitating the uptake of xenobiotics into the liver. The A allele was found associated with increased *OCT3* expression [147] which indicates increased hepatic

uptake and decreased efficacy as a result of reduced plasma levels. Our finding could be linked with this proposed mechanism. The depletion of the variant allele seen among lansoprazole users could be explained by reduction of its therapeutic efficacy. However, it is unknown whether lansoprazole is an OCT3 substrate. In addition, the small effect size seen with this association lessens its clinical importance.

5.1.2.2 Bendroflumethiazide-rs3743527 (ABCC1)

The T allele at rs3743527 (C>T) SNP in ABCC1 transporter was also seen underrepresented among bendroflumethiazide users. However, it is unknown if bendroflumethiazide is an ABCC1 substrate. It has been suggested that the majority of the drug is eliminated by a non-renal route as the drug seems to be extensively metabolized in the liver [148]. ABCC1 is expressed in the basolateral hepatic membrane facilitating the efflux of xenobiotics into the blood. SN-38, the active irinotecan metabolite, is transported by hepatic ABCC1 to increase circulating blood concentrations. The rs3743527 (C>T) polymorphism has been previously linked with decreased irinotecan-induced neutropenia [149] suggesting that this would be a loss-of-function variant. Nevertheless, it is unclear how carrying this variant could be linked with decreased tolerability to bendroflumethiazide, as shown in our findings. However, if the drug was found to be a substrate for ABCC1, then decreased hepatic efflux into blood could result in reduced plasma levels and efficacy. On the other hand, decreased ABCC1 function in the kidney could result in increasing systemic exposure and toxicity of the drug.

5.1.2.3 Gabapentin-rs8187843 (ABCC1)

Another *ABCC1* mutation, rs8187843 (G>A), was found to be associated with decreased frequency of the variant allele in the gabapentin treated group. It is unknown if gabapentin is a substrate for ABCC1 transporter. However, interestingly, this transporter is expressed in the basolateral membrane of the blood-brain barrier (BBB) protecting the brain from the entry of foreign substances; and therefore, genetic variability in this transporter could affect the pharmacokinetics of central nervous system (CNS) drugs which are substrates for the transporter. Reduced and increased the transporter function could increase and decrease the cerebral concentration of these agents, respectively.

Our research team is collaborating with Professor Kathy Giacomini (world leader in drug transporters) to establish whether or not the above-mentioned drugs lansoprazole, bendroflumethiazide, and gabapentin are transported by OCT3, ABCC1, and ABCC1 transporters respectively.

5.1.2.4 Rosuvastatin-rs2231135 (ABCG2)

The rs2231135 (A>G) ABCG2 transporter variant was overrepresented among rosuvastatin users suggesting increased tolerability to the drug. The majority of the drug, 72%-90%, is eliminated via the hepatic route [150] suggesting the importance of hepatic efflux transporters in its elimination. Rosuvastatin is a substrate for the ABCG2 efflux transporter [150]. The G allele at rs2231135 (A>G) *ABCG2* gene was associated with increased methotrexate (an ABCG2 substrate) mucositis side effects [151] which could be attributable to increased methotrexate plasma levels as a result of reduced ABCG2 efflux activity. However, the decreased activity of this transporter could increase rosuvastatin therapeutic efficacy due to increased residence time in the liver, which is the primary site for rosuvastatin pharmacological action. Our finding is consistent with this mechanism. In fact, our finding is further supported by multiple previous pharmacogenomic studies in which another independent *ABCG2* variant, rs2231142 (G>T), has been consistently linked with increased rosuvastatin efficacy as indicated by increased low-density lipoprotein (LDL) reduction [152].

5.2 Associations replicating previous studies findings

5.2.1 The combined cohort results (n = 10).

5.2.1.1 Gliclazide-rs1057910 (CYP2C9)

The antidiabetic drug gliclazide is primarily metabolized by CYP2C9 and CYP2C19 enzymes [153]. The reduced activity variant *CYP2C9*3* (rs1057910 (A>C)) has been shown to be correlated with increased risk of gliclazide-induced hypoglycemia [129]. Consequently, prescribers could reduce the daily dose to overcome this side effect, an effect consistent with that seen in the *CYP2C9*3* carriers in the Scottish combined data.

5.2.1.2 Simvastatin-rs2231142 (ABCG2)

The T allele at rs2231142 variant in the *ABCG2* gene is associated with a large decrease of simvastatin clearance [130]. The ABCG2 transporter is expressed in the apical membrane of the liver, facilitating drug efflux and excretion. The decreased clearance observed with carriers of this variant indicates reduced ABCG2 function. This could be associated with increased simvastatin hepatic concentration and, in turn, increased efficacy as the liver is where simvastatin works to exert its therapeutic efficacy. This is consistent with our observation that this variant allele was associated with being less likely to stop the treatment; probably due to increased efficacy.

5.2.1.3 Methotrexate-rs9895420 (ABCC3)/rs1128503 (ABCB1)

Methotrexate is mainly eliminated via renal excretion, and it is an ABCC3 and ABCB1 substrate [154]. ABCC3 is expressed in the apical renal membrane transporting its substrates into the urine. The rs9895420 (T>A) mutation in this transporter has been shown to be associated with increased ABCC3 activity, lower

methotrexate efficacy, and lower risk of thrombocytopenia [131]. Consistent with this, I have also shown that carriers of this variant allele were less likely to stop their methotrexate treatment. Thrombocytopenia can be a life-threatening side effect, and therefore, the advantage of reduced toxicity with methotrexate can outweigh the disadvantage of reduced efficacy especially given the fact that methotrexate is usually combined with other disease-modifying antirheumatic drugs (DMARDs) to improve therapeutic efficacy. This would explain our findings that carriers of the rs9895420 (T>A) had lower odds of stopping methotrexate treatment. In contrast, the reduced activity allele (A) at rs1128503 (A>G) SNP in ABCB1 were linked to increased methotrexate side effects [132]. This could explain our finding that this allele is associated with increased risk of stopping methotrexate treatment. It is expected that these observations are attributable to decreased renal ABCB1 efflux function resulting in increased methotrexate plasma concentration and toxicity.

5.2.1.4 Carbamazepine-rs762551(CYP1A2)/ rs2242480(CYP3A4)/ rs1128503(ABCB1) (3 associations)

The C allele at rs762551 (C(*minor allele*)>A) *CYP1A2* and the T allele at rs2242480 (C>T) *CYP3A4* have been linked with increased and decreased carbamazepine plasma concentrations respectively [133,134]. In our data, the two variants were associated with being more or less likely to decrease carbamazepine dose, respectively. These two clinical behaviours are consistent with the changes in the plasma concentrations observed in the two studies outlined above. In addition, the A allele at rs1128503 (A(*minor allele*)>G) ABCB1 SNP, has been connected with increased carbamazepine clearance in African Americans but not Caucasian population [135]. The authors of this study were not able to determine why there was a difference in the results between these two ethnic groups. However, increased carbamazepine clearance seen

with African Americans could result in decreased carbamazepine efficacy which is consistent with our finding in the British cohort that carriers of the minor allele (A) were more likely to stop the treatment

5.2.1.5 Pioglitazone-rs10509681 (CYP2C8)

It has been reported that Caucasian users of pioglitazone (a CYP2C8 substrate) harbouring the *CYP2C8*3* (rs10509681 (T>C)) variant experience lower pioglitazone plasma levels as compared to non-carriers [136]. Consistently, and probably due to decreased efficacy, I have observed carriers of this variant allele were more likely to stop the treatment.

5.2.1.6 Warfarin-rs2242480 (CYP3A4)

I identified an association between stopping warfarin treatment (a CYP3A4 substrate) and the rs2242480 (C>T) polymorphism in CYP3A4 enzyme. This is in consistent with a previous finding that the minor allele (T) was correlated with increased warfarin clearance [137], which could result in reduced efficacy and consequent treatment cessation.

5.2.1.7 Atorvastatin-rs2032582 (ABCB1)

The minor allele (A) at rs2032582 (A>T) SNP in ABCB1 is associated with enhanced atorvastatin efficacy as indicated by increased LDL reduction [138]. Our results are consistent with this, with the A allele being associated with a lower likelihood of treatment cessation. Of note, this is the third association I see confirming that reduced ABCB1/ABCG2 function is correlated with enhanced statin efficacy.

5.2.1.8 Clopidogrel-rs1057910 (CYP2C9)

Clopidogrel is a prodrug which requires metabolic activation to exert its therapeutic efficacy. This activation is catalyzed by CY3A4, CYP1A2, CYP2B6, CYP2C19, and CYP2C9 enzymes [155]. The loss-of-function variant *CYP2C9*3* has been reported to decrease clopidogrel efficacy [139]. Decreased efficacy could be also associated with decreased side effects. Consistently, I have identified that carriers of this variant allele are less likely to decrease clopidogrel dose. This can be attributable to decreased activation and, consequently, decreased side effects which may manifest as dyspepsia or even haemorrhage.

5.2.2 The UKBB (cross-sectional) cohort results (n=8).

5.2.2.1 Amitriptyline-rs1065852 (CYP2D6)

Amitriptyline (a tricyclic antidepressant (TCA)) is mainly metabolized by CYP2C19 into another active TCA nortriptyline. Nortriptyline is then metabolized primarily by CYP2D6 into less active metabolites (10-hydroxy metabolites) [156]. The reduced activity variant *CYP2D6*10* (rs1065852 (G>A)) has been shown to be associated with increased nortriptyline plasma levels [140]. This might be correlated with increased side effects. Consistent with this, our results from the UKBB cohort show that this variant was underrepresented among amitriptyline users compared to non-users suggesting decreased tolerability to the treatment.

5.2.2.2 Amlodipine-rs1045642 (ABCB1)

Our results also show a modest reduction in the distribution of the G allele at rs1045642 (A>G) in ABCB1 among amlodipine users suggesting reduced tolerability to the drug. In concordance with this result, this variant allele has been reported to be connected with increased amlodipine plasma levels [141] which could be associated with increased side effects.

5.2.2.3 Simvastatin-rs4149056 (SLCO1B1)

I have also identified that the well-known reduced function SNP *SLCO1B1*5* (rs4149056 (T>C)) is significantly lower within users of simvastatin (a SLCO1B1 substrate). This is in alignment with the well-established fact that this variant has been frequently seen to be connected with increased simvastatin-induced myopathy [142]. SLCO1B1 facilitates simvastatin uptake into the liver, and as a result of its reduced activity, simvastatin plasma levels are increased, which increases the risk of myopathy.

5.2.2.4 Clopidogrel-rs12248560/rs4244285 (CYP2C19) (2 associations)

As mentioned earlier, clopidogrel is activated by multiple CYP enzymes, including CYP2C19. The two well-known variants *CYP2C19*17* (rs12248560 (C>T)) and *CYP2C9*2* (rs4244285 (G>A)), defining the ultra-rapid and poor metabolizer phenotypes respectively, have a well-established correlation with increased and decreased clopidogrel efficacy as indicated by decreased and increased cardiovascular events respectively [143,144]. Consistent with these known associations, I have found that the distribution of the variant alleles of these two SNPs is higher and lower among clopidogrel users from the UKBB cohort compared to non-users, respectively. The higher distribution of the *CYP2C19*17* variant indicates increased tolerability

which is most likely due to increased efficacy and vice versa with the *CYP2C9*2* variant. Clopidogrel efficacy could be evaluated by observing the frequency of cardiovascular (CV) events in a specific group of patients. For example, if carrying a certain variant results in increased clopidogrel efficacy as indicated by reduced CV events, then it is expected that carriers of this variant would be overrepresented among clopidogrel users.

5.2.2.5 Citalopram-rs28371725 (CYP2D6)

The *CYP2D6*41* (rs28371725 (C>T) variant, which is associated with decreased enzyme activity, has been reported to be linked with increased citalopram (a CYP2D6 substrate) efficacy [145]. Our findings also show a modest increase in the distribution of the minor allele (T) within citalopram users from the UKBB cohort, suggesting increased tolerability in term of efficacy.

5.2.2.6 Pioglitazone-rs10509681 (CYP2C8)

In the combined cohort results above, I have shown that the *CYP2C8*3* variant was associated with higher odds to stop pioglitazone treatment; a result which was supported by a previous finding. Interestingly, I was also able to see a similar trend of association among pioglitazone users from the UKBB. This variant allele was less distributed within this treatment group suggesting decreased tolerability to the drug.

5.2.2.7 Fluoxetine-rs1065852 (CYP2D6)

I have also shown that the *CYP2D6*10* variant is underrepresented within fluoxetine (a CYP2D6 substrate) users from the UKBB cohort. This could be due to increased toxicity as the variant has shown a correlation with increased fluoxetine plasma concentrations from a previous study [146].

5.3 Potential important associations which didn't pass Bonferroni significance level.

5.3.1 The combined cohort results (n = 12).

5.3.1.1 Clopidogrel-rs12353214 (PTGS1(COX-1))

As this is the closest association to Bonferroni significance level occurring in an interesting gene, I decided to examine this association for replication from the UKBB primary care data. This association, along with its replication result, will be discussed in detail in chapter V.

5.3.1.2 Valsartan-rs4918758 (CYP2C9)

Among all of the other important CYP enzymes, CYP2C9 has been recognized as the only CYP enzyme responsible for metabolizing valsartan [157]. The liver has been identified as the main elimination route, but only 20% of the dose is subject to hepatic metabolism. Even though this is a minimal contribution of hepatic metabolism in the total elimination process, I have identified the rs4918758 (T>C) variant in CYP2C9 to be associated with a lower likelihood of a decrease in valsartan dose. Interestingly, here I see this variant showing a similar direction of effect with the same phenotype as has been identified with quinine treatment, which was one of top findings from the combined Scottish data that was replicated in UK Biobank primary care data (presented and discussed in chapter V).

5.3.1.3 Ezetimibe-rs3842 (ABCB1)

In a previous study [158], induction of the intestinal ABCB1 expression by rifampicin resulted in a marked reduction in the ezetimibe plasma concentration and efficacy. Our study reveals that the variant allele (C) at rs3842(T>C) SNP in the ABCB1 transporter is associated with 89% increased risk of stopping ezetimibe treatment.

This variant has been previously reported to be correlated with increased efavirenz systemic exposure [159] suggesting reduced ABCB1 transportation activity. Similarly, this variant could increase ezetimibe plasma concentration and side effects (e.g. fatigue, abdominal and muscle pain) resulting in stopping the drug, which is what we have found.

5.3.1.4 Enalapril-rs2244614 (CES1)

The prodrug enalapril has been shown to be readily hydrolyzed by carboxylesterase1 (CES1) into its active metabolite enalaprilat [160]. This enzyme also contributes to the first step in the activation of capecitabine into its active metabolite 5-fluorouracil [161]. The minor allele (G) at rs2244614 (G>A) SNP in *CES1* gene has been associated with increased CES1 overexpression and increased capecitabine-induced toxicity [161]. However, our results show that this variant allele is associated with decreased odds of dose reduction of enalapril suggesting reduced toxicity which is inconsistent with the direction of effect seen with capecitabine treatment.

5.3.1.5 Lansoprazole- rs9282564 (ABCB1)

Lansoprazole is a substrate for the ABCB1 transporter [162]. I have noted that the rs9282564 (T>C) in the *ABCB1* gene is associated with a lower chance of decreased lansoprazole daily dose. This variant was previously linked with increasing plasma concentrations for multiple ABCB1 substrates [163]. However, it is unclear how this variant is associated with our observed phenotype.

6.3.1.6 Atenolol-rs628031 (SLC22A1)

It has been recently discovered that atenolol is transported by the OCT1 transporter expressed in the apical intestinal membrane [164] suggesting the potential role of this transporter in affecting atenolol absorption. Here, we report an association between

the minor allele (A) at rs628031(A>G) SNP in SLC22A1 and increased odds of stopping atenolol treatment. Of note, a similar trend of association was also detected from the UKBB cross-sectional data with carriers of this variant found to be depleted among atenolol group (OR = 0.97, p= 0.0245). This could be related to an increased intestinal accumulation of atenolol, causing abdominal side effects, as well as decreased bioavailability and efficacy. Interestingly, the same minor allele (A) was previously reported to be linked with increased metformin (a known OCT1 substrate) gastrointestinal tract toxicity [165] providing more support for our proposed mechanism of drug-gene interaction with atenolol.

5.3.1.7 Digoxin/nifedipine-rs4728709 (ABCB1) (2 associations)

Digoxin and nifedipine are both ABCB1 substrates [166,167]. I found that carriers of the A allele at rs4728709 (G>A) in ABCB1 transporter were less likely to decrease digoxin dose and less likely to stop nifedipine treatment. These phenotypes suggest increased tolerability and reduced side effects. Consistent with these findings, this variant was reported to be correlated with decreased vincristine-induced neurotoxicity [168]. All of these observations indicate that this variant could be linked with increased ABCB1 efflux activity.

5.3.1.8 Ranitidine-rs316019 (SLC22A2)

Ranitidine has been recently identified as an OCT2 substrate [169]. OCT2 is expressed in the basolateral renal membrane facilitating drug uptake into the kidney. I found that carriers of the minor allele (A) at rs316019 (A>C) variant in SLC22A2 transporter were less likely to decrease their ranitidine dose, suggesting decreased side effects. This allele has been reported to affect multiple drugs but with conflicting directions of effects. For example, it has been correlated with lower cisplatin/anthracyclines toxicity (in concordance with our finding with ranitidine), but it has also been found linked with decreased metformin/l-tryptophan clearance [170].

5.3.1.9 Valproic acid-rs2279343 (CYP2B6)

Valproic acid is mainly metabolized by CYP2A6, CYP2B6, and CYP2C9 enzymes [171]. Our results show that the *CYP2B6*4* (rs2279343 (A>G)) variant is correlated with an increased chance to decrease the valproic acid dose, suggesting increased side effects risk. However, the *CYP2B6*4* variant is a gain-of-function variant [172] and therefore the mechanism by which our finding occurred is not clear.

5.3.1.10 Clopidogrel-rs3815583 (*CES1*)

Around 85% of clopidogrel is converted into inactive metabolites via carboxylestrase1 (CES1) [173]. I found that the variant allele (C) at rs3815583 (A>C) SNP in the *CES1* gene is associated with a lower tendency to decrease clopidogrel dose which could indicate decreased risk for adverse events. In concordance with this result, the C-allele has been reported to be associated with decreased isoniazid (a suggested CES1 substrate)-induced hepatotoxicity [174].

5.3.1.11 Simvastatin-rs1080985 (*CYP2D6*)

In previous studies [175], reduced function *CYP2D6* variants were correlated with increased simvastatin (a CYP2D6 substrate) efficacy and side effects. On the other hand, I show that the C-allele at rs1080985 (G>C) *CYP2D6* is associated with a lower tendency toward decreased simvastatin dose, suggesting reduced side effects. This would require rs1080985 to be a gain-of-function variant in CYP2D6. Interestingly, this is supported by a previous study in which the C-allele was found to be associated with increased debrisoquine (a CYP2D6 substrate) metabolism [176].

5.3.2 The UKBB (cross-sectional) cohort results (n=8).

5.3.2.1 Atorvastatin-rs17287570 (ABCC1)

It has been shown that atorvastatin is an ABCC1 substrate [177]. The C-allele at rs17287570 (A>C) in the ABCC1 transporter has been previously associated with increased irinotecan-induced myelosuppression [178]. However, the authors mentioned that this association was significant in their first study but not in the second study and after combining both cohorts. SN-38 is the active metabolite of irinotecan, catalyzed by liver esterase, and is transported by the ABCC1 transporter. This transporter is expressed in the basolateral hepatic membrane pumping its substrates into the blood. Therefore, the increased activity of this transporter could be associated with irinotecan-induced side effects suggesting that the C-allele could be an increased activity variant. Our findings show that the C-allele at this *ABCC1* SNP was underrepresented among atorvastatin users. Increased ABCC1 hepatic activity could decrease atorvastatin hepatic concentration (i.e., decreased efficacy) and increase its systemic exposure (i.e. increased side effects).

5.3.2.2 Esomeprazole-rs12248560(CYP2C19)/rs1128503/rs1045642(ABCB1) (3 associations)

Esomeprazole is primarily metabolized by the CYP2C19 enzyme, and it is a substrate for the ABCB1 transporter [179]. Genetic variability in *CYP2C19* was shown to affect esomeprazole efficacy. For instance, carriers of *CYP2C19*2/*3* reduced activity variants have decreased esomeprazole metabolism and better response [180,181]. Here, I have also shown that patients carrying the CYP2C19*17 (rs12248560) variant (ultra-rapid metabolizers [182]) are enriched within esomeprazole treatment group in UKBB. However, extensive metabolism will result in decreased efficacy and, therefore, one would predict an under-representation of

this SNP in esomeprazole users. I have also noted the same phenotype (i.e. overrepresentation of the allele) occurring with carriers of the A allele at rs1128503 (A>G) *ABCB1* SNP which could be attributable to better response due to reduced ABCB1 efflux activity as what has been observed with other multiple drugs [183]. The other *ABCB1* SNP, rs1045642 (A>G), was extensively studied, but it shows different directions of effects with different drugs [184]. Our findings suggest that harbouring this variant could be associated with decreased tolerability for esomeprazole. Overall, these findings from the UKBB are not clear.

5.3.2.3 Methotrexate-rs4793665 (ABCC3)

I have shown a correlation between methotrexate and the rs9895420 (T>A) *ABCC3* SNP from the combined Scottish cohort, as discussed above. Here, in UKBB I have also detected a very modest effect of the minor allele (C) at rs4793665 (C>G) SNP in the ABCC3 transporter on methotrexate use with this variant being underrepresented in this treatment group. This minor allele has been previously correlated with increased hepatic efflux of morphine metabolites into plasma [185].

5.3.2.4 Allopurinol-rs2231135 (ABCG2)

In an association replicated multiple times from previous work, carriage of the reduced function T-allele at rs2231142 (G>T) in the ABCG2 transporter was associated with poor allopurinol response and increased likelihood for dose increase [186]. I couldn't find this SNP to be associated with any phenotype I have used in this study except for the strong association seen among allopurinol users from the UKBB cohort. However, this was clearly driven by disease (gout) rather than drug response. ABCG2 is involved in the transportation of uric acid and it has been shown that decreased intestinal ABCG2 function is associated with increased systemic exposure of uric acid [187]. Intestinal ABCG2 transporter returns substrates back

into the intestinal lumen and therefore, reducing this function could result in increasing serum uric acid levels and increased risk of gout. Nevertheless, I have identified another disease-independent SNP in the same gene (ABCG2), rs2231135 (A>G) that is underrepresented among allopurinol users from the UKBB cohort. It is unclear why this SNP seemed to affect allopurinol response only while being gout disease independent, but genotype dosage differences between different variants could play a role. This SNP has been previously shown to be correlated with increased methotrexate (an ABCG2 substrate)-induced mucositis indicating reduced ABCG2 function [188]. Our findings regarding allopurinol support previous results that decreased ABCG2 activity could be linked with poor allopurinol response as indicated by the low distribution of the rs2231135 (A>G) SNP among allopurinol users but the exact mechanism of this DGI is unknown.

5.3.2.5 Enalapril-rs7317112 (ABCC4)

Enalapril is a prodrug which undergoes hepatic hydrolysis by esterases to produce its active metabolite enalaprilat. It has been recently reported that hepatic ABCC4 transporters are responsible for the translocation of enalaprilat from hepatocytes into the systemic circulation for the drug to exert its blood pressure-lowering efficacy [189]. In the same context, our results reveal an *ABCC4* polymorphism, rs7317112 (A>G), to be overrepresented in the enalapril treatment group indicating increased tolerability to the drug. This can be explained by an increase in ABCC4 function, which could result in increased systemic enalaprilat exposure and efficacy. In fact, the increased activity could be predicted from a previous study in which the minor allele (G) has been correlated with decreased methotrexate-related mucositis [190].

5.3.2.6 Clopidogrel-rs9332197 (CYP2C9)

As explained earlier, clopidogrel is activated by multiple CYP enzymes, including CYP2C9. I have also discussed the link between the *CYP2C9*2* reduced function variant and loss of clopidogrel efficacy. In UKBB, I also detected another independent *CYP2C9* SNP, rs9332197 (T>C), which has a similar effect as indicated by the reduced distribution of the variant allele (C) among clopidogrel users. This variant has also been reported to affect warfarin efficacy [191].

6. Conclusion

The present work has provided for the first time a large-scale coverage of clinical pharmacogenomic associations between 162 genetic variants in important enzymes and transporters and 50 commonly used chronic drugs using 3 different drug response phenotypes (drug-stop, dose-decrease, and changes in genotype distribution among drugs' users) in two cohorts (the Scottish cohort + the UKBB cohort). I have generated a total of 24,300 drug-genetic variants associations results which are accessible online via this link: https://c1abo933.caspio.com/dp/d81f7000c3c1854d29104a49b1d8. I have identified a total of 12 novel drug-genetic variant associations (passed Bonferroni significance level). I have also identified 19 results replicating previous study findings. Finally, I also show a total of 20 potentially important novel associations occurring in relevant genes. However, it is important to recognise that results that didn't pass Bonferroni significance should be interpreted with caution. We included these results in our discussion only to shed some light on candidate results which could merit further investigation in the future but not to claim definitive conclusions from them.

We acknowledge that drug response phenotype defined from the UKBB cross-sectional data approach can be easily influenced by disease associations, and generally difficult to interpret. In fact, we used this approach because the UKBB cross-sectional data was the only cohort available to us at the beginning of this PhD. The complete longitudinal Scottish cohort data didn't become available to us until after around two years from starting this PhD. Once these cohorts become available, we used them to define more accurate phenotypes (i.e. drug-stop and dose-decrease). In the next chapter (drug-drug-gene interactions), we used a similar methodology and phenotype definitions as used in this chapter.

Chapter IV:

Drug-Drug-Gene Interactions for the Most Commonly Used Chronic Drug Combinations in the UK

Abstract

Most of the current work in relation to drug-interactions focuses on drug-drug interactions and, to a lesser extent, drug-gene interactions. In contrast, there is very limited research on drug-drug-gene interactions (DDGIs). In the present study, I attempted to uncover novel and clinically important drug-drug-gene interactions among commonly used chronic drug combinations in the UK. I studied the associations between 50 common chronic drug combinations and 162 selected genetic variants in important enzymes and transporters. These associations were studied in two cohorts: a combined cohort of longitudinal prescribing data for 3 Scottish cohorts and the UKBB cross-sectional prescribing data. In the combined cohort, these drug-drug-variant combinations have been studied with respect to a 'drug-stop' or 'dose-decrease' phenotype while in the UKBB cross-sectional data we studied differences in the genotype distribution between drug combinations of users compared to non-users. I have identified 3 and 7 novel drug-drug-gene associations which were significant after Bonferroni correction from the combined Scottish cohort and the UKBB cohorts respectively.

From the combined cohort, users of the bisoprolol-atorvastatin combination who carry the rs9516519 (T>G) *ABCC4* variant and users of the simvastatin-metformin combination who carry the rs622342 (C(*minor allele*)>A) *SLC22A1* minor allele were more likely to stop one of the two drugs during their combined use per allele (ORs = 4.72 (2.44-9.13) and 3.1 (1.71-5.62), p-values = 1.48×10^{-5} and 9.3×10^{-5} respectively). In addition, the G allele at rs1967120 (G(*minor allele*)>A) *ABCC1* SNP has been observed to be correlated with increased risk of dose reduction of either metformin or gliclazide when taken concomitantly (OR = 1.21 (1.1-1.33), p=9.3×10^{-5}).

The results from the UKBB cohort show depletion of the minor allele for the following four associations :**1)** bendroflumethiazide-metformin-rs2199939 (C>T) *ABCG2* SNP (OR = 0.81 (0.73-0.9); p=5.29 × 10^{-5}) **2)** atorvastatin-metformin-rs2293001 (C>T) *ABCC5* SNP (OR = 0.89 (0.83-0.94); p= 1.6 × 10^{-4}) , **3)** atorvastatin-metformin-rs17731538 (G>A) *ABCG2* SNP (OR = 0.85 (0.78-0.93); p=1.6 × 10^{-4}) , and **4)** simvastatin-metformin-rs215095 (G(*minor allele*)>A) *ABCC1* SNP (OR = 0.9 (0.84-0.95); p=2.606 × 10^{-4})

In contrast, the distributions of the minor allele was enriched in the following 3 associations: **1)** simvastatin-furosemide-rs4148739 (T>C) *ABCB1* SNP (OR = 1.26 (1.13-1.41); p=6.07 × 10^{-5}), **2)** atorvastatin-metformin-rs10937158 (T(*minor allele*)>C) *ABCC5* SNP (OR = 1.13 (1.06-1.2); p=1.04 × 10^{-4}) , and **3)** amlodipine-atorvastatin-rs3735451 (T>C) *CYP3A4* SNP (OR = 1.21 (1.1-1.34); p=1.3 × 10^{-4}) .

1. Introduction

To date, the literature on drug-drug-gene interactions (DDGIs) is limited. In my review, which has been presented in chapter I, only 33 DDGI studies and case reports were identified (17 with metabolizing enzyme genes and 16 with transporter genes). The few papers on drug-drug-metabolizing enzyme genes focused only on *CYP2C9*, *CYP2C19*, and *CYP2D6* genes and only for a few drugs. Genetic variants in other important genes such as *CYP3A4, CYP2C8, CYP2E1, CYP1A2* and *CYP2B6* were not detected in my search for DDGIs reports. The gap becomes even larger when we consider drug-drug-transporter gene interactions. Transporters are generally poorly studied compared to metabolizing enzymes, for drug interactions. This is probably due to the only recent evolution of transporter science after it was thought that metabolizing enzymes were the only major contributors to drug elimination. I have identified a limited number of DDGI papers which have studied limited drugs and limited variants in the ABCB1, SLCO1B1, and OCT1 transporters. In addition, there were no publications for DDGIs involving the ABCC1/2/3/4, SLCO2B1/1B3, OCT2/3, OAT1/2/3, and ABCG2 transporters. These kinds of interactions could intensify or ameliorate the predicted clinical outcomes seen with drug-drug only or drug-gene only interactions.

Based on this literature, it is clear that there is currently a large information gap regarding DDGIs. Therefore, in this study, we attempt to broadly cover this topic by considering a large variety of common drug combinations and studying their associations with our previously selected genetic variants in metabolizing enzymes and transporters.

2. Study Populations

Four UK cohorts have been the target of this study: 3 Scottish cohorts and one from across the UK. The Scottish data represents 3 cohorts: GoDARTs, Generation Scotland (GS), and GoSHARE longitudinal prescribing data which are combined together. The combined cohort, along with the UKBB cross-sectional data, are used for this study. For a detailed description of these cohorts, see chapter II.

3. Methodology

The methods we have applied can be summarized in the below five steps:

3.1 Selection of candidate common chronic drug combinations

3.1.1 Identifying a list of common drug combinations in the UK.

In chapter III, I utilized the UKBB cross-sectional prescribing data to produce a list of the top 122 most frequently used drugs in the UK. This list was used to generate a second list of the most frequently prescribed drug combinations. A drug combination is considered common when both drugs are within our list of the 122 top drugs. Therefore, a matrix of size 122*122 was produced, and the frequency of each possible unique combination (n=7333) was calculated.

Table 7 below clarifies the structure of the matrix produced.

	No.	1	2	3	4	…	…	122
No.	Drug	Paracetamol	Aspirin	Ibuprofen	Simvastatin	…	…	…
1	Paracetamol	0	14231	28472	10669	…	…	…
2	Aspirin	NA	0	7261	27022	…	…	…
3	Ibuprofen	NA	NA	0	4584	…	…	…
4	Simvastatin	NA	NA	NA	0	…	…	…
…	…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…	…
122	…	…	…	…	…	…	…	…

Table 7: Part of the matrix which consists of a head row of our top 122 common drugs and a first column of the same list of drugs. This enabled us to identify and calculate the frequency of each possible unique combination between the drugs in the list. The frequencies are recorded in the cells for each combination. For example, the frequency of using paracetamol-aspirin combination is 14231 times. Combinations where both drugs are the same take the frequency of "0" and duplicated combinations where the frequency has already been calculated before take the frequency of "NA". The frequencies of only unique combinations with two unique drugs appear in the matrix (n=7333).

This matrix was then converted into a table showing the frequency of unique drug combinations starting from the highest usage frequency.

3.1.2 Development of a customized DDIs database for the 7333 combinations

Table 8 below summarizes the steps for producing the drug-drug interactions database showing 3 different scenarios for interactions.

Pharmacokinetic drug interactions can be inhibitory or induction interactions depending on whether the metabolizing enzymes (MEs) and/or transporters (Trs) of the victim drug are inhibited or induced by the perpetrator drug. The majority of DDIs are expected to be inhibitory rather than induction interactions as the majority of perpetrator drugs are inhibitors rather than inducers. Therefore, we will focus on identifying inhibitory drug interactions occurring in the 7333 drug combinations. To

do so; first, we extracted information about the MEs/Trs involved in PK of the 122 commonly used drugs in UK Biobank from the DrugBank database (described in chapter II). MEs/Trs inhibited by each drug were also extracted, resulting in each drug having two columns of information: "substrate of" and "inhibitor of". I then extracted information from this table into the table of 7333 combinations, so each drug in each combination has the two columns of information. This table was then utilized to identify possible routes of interaction between drugs in each combination. There are two possible mechanisms for interaction within each drug combination: drug1 as the substrate drug and drug 2 as an inhibitor drug or vice versa. Considering both sides of the interaction, I produced results showing the interacting genes for each combination.

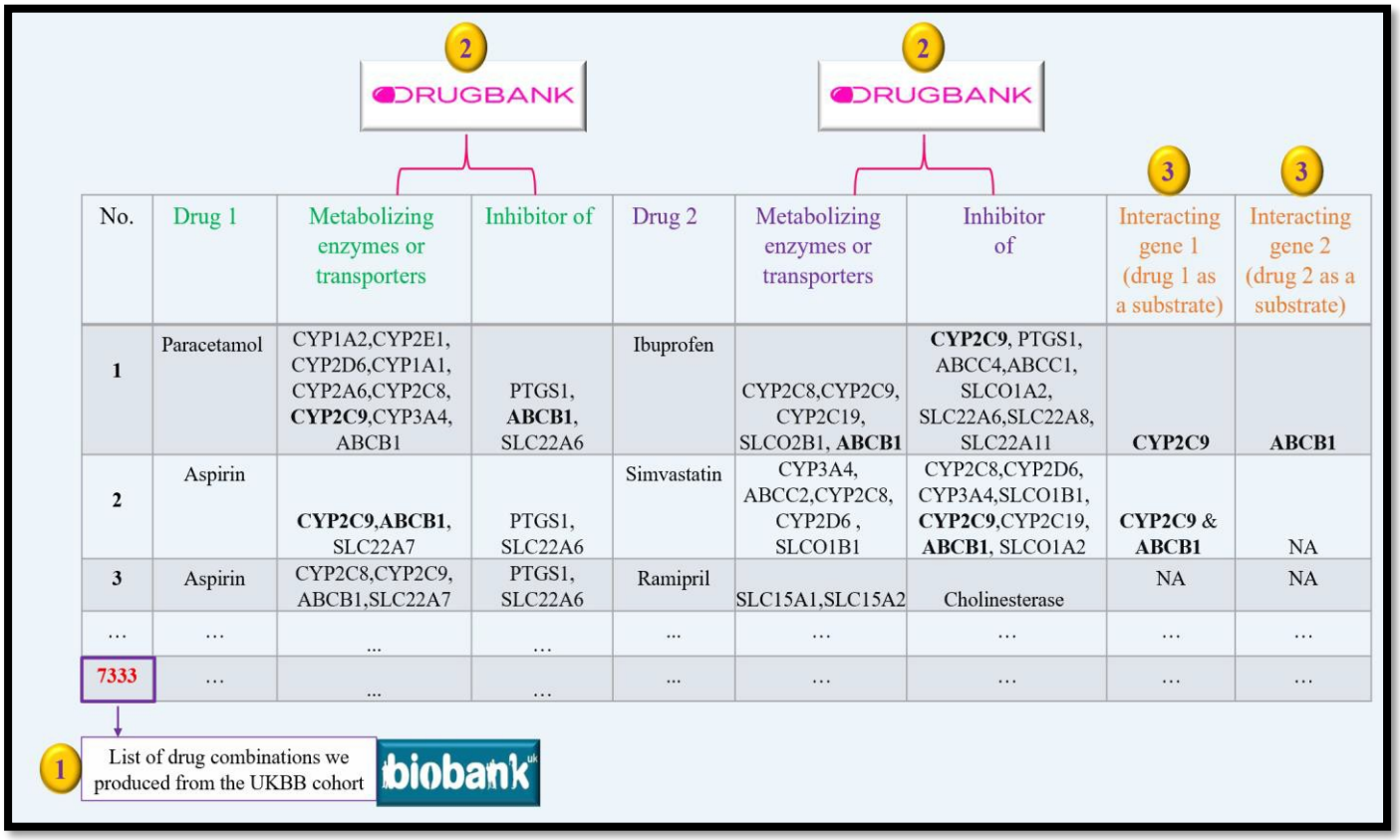


Table 8: The steps of producing our drug-drug interactions database for common drug combinations in the UK showing 3 different scenarios of interactions. We first produced a list of 7333 different drug combinations used in the UK from the UKBB cohort **(1).** Then, information about metabolizing enzymes or transporters and enzymes or transporters inhibited by each drug were extracted from DrugBank for each drug in each combination **(2/2)**. Finally, the interacting gene was recognized for each combination considering both modes of interaction **(3/3).**

The table shows 3 possible scenarios of interaction. The first row shows an interaction occurring where both drug 1 and drug 2 are the victim drug (e.g. ibuprofen is an inhibitor for CYP2C9 which is one of the metabolizing enzymes for paracetamol. On the other hand, paracetamol is an inhibitor for ABCB1 which is a transporter for ibuprofen). The second row shows an interaction where only one of the two drugs is the victim drug while the third row shows a drug combination with an unknown interaction route.

3.1.3 Selecting common drug combinations of interest

Our results show that the 7333 combinations can be classified into two categories: interacting combinations (n = 1,503) and combinations with unknow routes of interactions (n = 5,830). The first category represents any combination where there is at least one route of interaction between the two drugs, while the second category shows combinations with no identifiable interaction between the two drugs. From each category, I selected the top 25 commonly used chronic combinations. Our final list of top 50 commonly used combinations (25 interacting combinations and 25 combinations with unknown routes of interactions) is shown in Table 9 below. As the knowledge regarding drug-metabolizing enzymes and transporters, and the ability of drugs to inhibit these enzymes or transporters, is limited we chose to include 25 combinations where there is as yet no known route of interaction between the drugs used in combination, in addition to the 25 where routes of interaction were better recognized.

		1) Top 25 interacting chronic drug combinations		
No.	UKBB Rank	Drug combination	Interacting gene (**Drug 1** as a substrate/ drug 2 as an inhibitor)	Interacting gene (**Drug 2** as a substrate/ drug 1 as an inhibitor)
1	4	Simvastatin-Amlodipine	*CYP3A4*	NA
2	6	Simvastatin-Omeprazole	*CYP3A4*	*CYP2C8, CYP3A4, CYP2C9, CYP2C19*
3	10	Simvastatin-Lansoprazole	*CYP3A4, CYP2C8*	*CYP2C8, CYP3A4, CYP2C19*
4	16	Simvastatin-Bisoprolol	NA	*CYP3A4*
5	19	Atenolol-Atorvastatin	*ABCB1*	NA
6	24	Omeprazole-Atorvastatin	*CYP2C8, CYP2C19, CYP2C9, CYP3A4*	*CYP3A4*
7	26	Simvastatin-Gliclazide	NA	*CYP2C9, CYP2C19*
8	28	Amlodipine-Atorvastatin	NA	*CYP3A4*
9	32	Omeprazole-Amlodipine	*CYP3A4 , CYP1A1*	NA
10	33	Simvastatin-Doxazosin	NA	*CYP2C19*
11	36	Simvastatin-Amitriptyline	*CYP2C8*	*CYP2C8, CYP2D6, CYP2C9, CYP2C19*
12	37	Omeprazole-Amitriptyline	*CYP2C8, CYP2C19*	*CYP2C19, CYP2C9, CYP1A2, CYP2D6*
13	38	Atorvastatin-Lansoprazole	*CYP3A4, CYP2C8*	*CYP3A4, CYP2C8, CYP2C19, ABCB1*
14	39	Simvastatin-Losartan	*CYP3A4, CYP2C8*	*CYP2C8, CYP3A4, CYP2C9*
15	40	Atorvastatin-Bisoprolol	NA	*CYP3A4*
16	41	Simvastatin-Clopidogrel	*CYP3A4, CYP2C8*	*CYP3A4, CYP2C9, CYP2C19*
17	42	Amlodipine-Lansoprazole	NA	*CYP3A4*
18	43	Atenolol-Lansoprazole	*ABCB1*	*ABCB1*
19	44	Simvastatin-Felodipine	*CYP3A4, CYP2C8*	*CYP3A4*
20	45	Simvastatin-Warfarin	*CYP2C8*	*CYP2C8, CYP3A4, CYP2C9, CYP2C19*
21	46	Simvastatin-Ranitidine	*CYP3A5, CYP3A4*	*CYP2D6, CYP2C19*
22	47	Simvastatin-Furosemide	*ABCC2*	NA
23	48	Simvastatin-Citalopram	NA	*CYP2D6, CYP3A4, CYP2C19*
24	49	Omeprazole-Bisoprolol	NA	*CYP3A4*
25	50	Atorvastatin-Clopidogrel	*CYP3A4, CYP2C8*	*CYP3A4, CYP2C9, CYP2C19, ABCB1*
		2) Top 25 chronic drug combinations with unknown interaction routes		
No.	UKBB.Rank	Drug combination	Interacting gene (**Drug 1** as a substrate/ drug 2 as an inhibitor)	Interacting gene (**Drug 2** as a substrate/ drug 1 as an inhibitor)
1	1	Simvastatin-Ramipril	NA	NA
2	2	Simvastatin-Bendroflumethiazide	NA	NA
3	3	Simvastatin-Metformin	NA	NA
4	5	Simvastatin-Atenolol	NA	NA
5	7	Simvastatin-Lisinopril	NA	NA
6	8	Bendroflumethiazide-Amlodipine	NA	NA
7	9	Bendroflumethiazide-Ramipril	NA	NA
8	11	Bendroflumethiazide-Atenolol	NA	NA
9	12	Ramipril-Amlodipine	NA	NA
10	13	Metformin-Gliclazide	NA	NA
11	14	Ramipril-Atorvastatin	NA	NA
12	15	Ramipril-Atenolol	NA	NA
13	17	Ramipril-Metformin	NA	NA
14	18	Bendroflumethiazide-Lisinopril	NA	NA
15	20	Amlodipine-Atenolol	NA	NA
16	21	Atorvastatin-Metformin	NA	NA
17	22	Simvastatin-Perindopril	NA	NA
18	23	Omeprazole-Bendroflumethiazide	NA	NA
19	25	Bendroflumethiazide-Atorvastatin	NA	NA
20	27	Omeprazole-Ramipril	NA	NA
21	29	Amlodipine-Metformin	NA	NA
22	30	Ramipril-Bisoprolol	NA	NA
23	31	Omeprazole-Atenolol	NA	NA
24	34	Amlodipine-Lisinopril	NA	NA
25	35	Bendroflumethiazide-Metformin	NA	NA

Table 9: List of top 50 commonly used drug combinations in the UK as classified into interacting combinations (n =25) and combinations with unknown interaction routes (n=25).

UKBB Rank = ranking of the frequency of use of the combination in the UK.

3.2 Selection of candidate genetic variants

In chapter III, I have shown the detailed process of selecting the genetic variants of interest in genes encoding drug-metabolizing enzymes and transporters. The same list of 162 SNPs in important enzymes and transporters selected previously will be considered here in this DDGI study.

3.3 Defining drug response phenotypes

I have also presented the detailed methods in defining drug response phenotypes in chapter III, considering both cross-sectional and longitudinal prescribing data. I briefly describe these in terms of DDGIs below.

3.3.1 In cross-sectional prescribing data (UK Biobank)

Here we again use deviation from HWE as an indication of selection seen where an allele is enriched or depleted in patients exposed to a particular drug combination. This enrichment or depletion may reflect one of two 'drug response' phenotypes:

- *Increased efficacy and/or decreased side effects:*

  this occurs when the variant allele is found to be enriched within the drug combination users compared to non-users. If those with a specific drug-drug-variant combination have a higher allele frequency, this would suggest that the enriched allele is beneficial for these patients in term of increased efficacy, low side effects, or both.
- *Decreased efficacy and/or increased side effects:*

  this occurs when the variant allele is found to be significantly depleted within the drug combination users compared to non-users. If those with a specific

drug-drug-variant combination have lower allele frequency; this would reflect that the depleted allele was disadvantageous for these patients in term of decreased efficacy, increased side effects, or both.

Of note, deviation from HWE can be caused by other reasons rather than using the drug combination such as a deviation occurring due to diseases for which the drug combination is used. We have dealt with this issue by using GeneAtlas database to identify whether the significant SNP we have discovered is correlated with diseases for which any of the two drugs are usually prescribed for. If a clear disease association was found, I exclude the result from our list of UKBB results. The second issue is that the discovered SNP might be correlated with using only one of the two drugs but not necessarily because of using the drug combination. To deal with this issue, I discuss our results in the light of our findings from our DGIs study from chapter III and explain whether the discovered SNP is exclusively correlated with using the drug combination only or it is also correlated with using one of the two drugs.

3.3.2 In longitudinal prescribing data (the combined cohort)

Two drug response phenotypes were studied: drug-stop and dose-decrease. After identifying patients who were prescribed the drug combination of interest, they were classified into cases and controls. Cases were those who stopped (i.e. had only one prescription) any of the two chronic drugs during their interaction. The control group represented patients who were on the same drug combination but had two or more prescriptions from both drugs during the interaction time.

Similarly, for the dose-decrease phenotype, individuals who reduced the daily dose of any of the two drugs during the interaction time were considered cases while those who never reduced their daily dose of any of the two drugs since being co-prescribed were considered controls.

3.4 Testing the association between the genetic variants and the phenotypes

Using "SNPassoc" R package, I ran a case-control genetic analysis under the log-additive model to explore the associations between the 162 selected genetic variants and each drug response phenotype for each cohort, for all 50 drug combinations. The Bonferroni adjusted p-value for significance was $p <= 0.00030$ (0.05/162).

3.5 Development of an online database to view the results

Given the large number of results produced, in parallel to work presented on DGIs in chapter III, I developed a dynamic, user-friendly online database to view all results visually (graphs) or as tables using Caspio software.

4. Results

For the combined Scottish data, examining the association between the drug-stop phenotype for 50 drug combinations and 162 genetic variants produced a total of 8,100 results; the same number of findings were also produced for the dose-decrease phenotype from the combined cohort taking the total findings into 16,200.

Considering the UKBB cross-sectional cohort, I tested changes in the genotype distribution of the 162 variants between users of each of the 50 drug combinations versus non-users which produced a total of 8,100 findings.

In order to facilitate viewing of the results, all results can be accessed via an online database under this link:

https://c1abo933.caspio.com/dp/d81f7000033513b3ab1c4431b5a2

There were 769 findings with a nominal significance level (p <= 0.05) from the combined cohort when considering both the drug-stop and dose-decrease phenotypes. 381 of these results belong to the 'interacting combinations' while the other 388 findings are from the 'combinations with unknown interaction routes' category.

The results from the UKBB cross-section cohort show a total of 489 findings with p-values <= 0.05. Of which, 239 results are from the 'interacting combinations', and the other 250 findings are from the 'combinations with unknown interaction routes' category.

The classification for the significant DDGI results differs from the DGI results classification presented in chapter III. The number of published DGIs papers (n ~ 23,000, according to PharmGKB) is much higher than the ones published on DDGIs (n = 33, according to our review in chapter I). This means it will not be possible to show prior published results that we replicate in our analysis, meaning the majority, if not all, of our DDGIs findings, should be considered novel but in need of replication. The results for our DDGI study are structured as follows: 1) the associations which have passed Bonferroni significance level; 2) potentially important associations from the 'interacting combination' category which didn't pass Bonferroni significance level; 3) potentially important associations which didn't pass Bonferroni significance level from the 'combinations with unknown interaction routes' category. In the below sections, I present results belonging to these 3 categories.

4.1 Significant associations after Bonferroni correction (p <=0.0003)

Table 10, at the end of the results section, summarizes all findings under this category.

4.1.1 The combined Scottish cohorts (n = 3).

There were 3 associations significant after Bonferroni correction in the combined Scottish cohorts; these are outlined below.

4.1.1.1 Atorvastatin-Bisoprolol-rs9516519 (ABCC4)

The *ABCC4* variant, rs9516519 (T>G) was strongly associated with atorvastatin-bisoprolol use ($p=1.48 \times 10^{-5}$). The G- allele was associated with a 4.72 (2.44-9.13) times increased risk of stopping any of the two drugs when administered together.

4.1.1.2 Simvastatin-Metformin-rs622342 (SLC22A1)

There was an interaction between simvastatin and metformin use in combination with rs622342 (C(*minor allele*)>A) in SLC22A1 ($p=9.3 \times 10^{-5}$). The C-allele was associated with a 3.1 (1.71-5.65) times greater odds of stopping any of the two drugs during their combined use.

4.1.1.3 Metformin-Gliclazide-rs1967120 (ABCC1)

We also found that the rs1967120 (G(*minor allele*)>A) variant in *ABCC1* gene is associated with increased odds of decrease in daily dose of either metformin or gliclazide when used in combination (OR 1.21 (1.01-1.33); $p= 9.3 \times 10^{-5}$).

4.1.2 The UKBB (cross-sectional) cohort results (n = 7)

There were 15 results significant after Bonferroni correction. However, using GeneAtlas it was apparent that 8 of these were associated with the underlying disease, with 7 appearing disease independent. These 7 results are outlined below:

4.1.2.1 Bendroflumethiazide-Metformin-rs2199939 (ABCG2)

Our results show a significant depletion of the T-allele at rs2199939 (C>T) in *ABCG2* gene within patients on this drug combination ($p=5.29 \times 10^{-5}$). Detection of the T-allele in this group is 19% (10%-27%) less likely compared to individuals not on this drug combination.

4.1.2.2 Simvastatin-Furosemide-rs4148739 (ABCB1)

There was a 26% (13%-41%) higher chance for the C-allele at rs4148739 (T>C) in *ABCB1* gene to occur among users of the simvastatin-furosemide combination compared to non-users ($p=6.07 \times 10^{-5}$).

4.1.2.3 Atorvastatin-Metformin-rs10937158/rs2293001 (ABCC5) / rs17731538(ABCG2) (3 associations)

Two *ABCC5* variants, rs10937158 (T(*minor allele*)>C) and rs2293001 (C>T) , were associated with atorvastatin-metformin use (p-values = 1.04×10^{-4} and 1.6×10^{-4} respectively). The odds carrying the T-allele was 13% (6%-20%) higher and 11% (6%-17%) lower among users of this drug combination for both variants respectively.

Another *ABCG2* variant, rs17731538 (G>A), was also associated with this drug pair as there was a 15% (7%-22%) lower odds for carriage of the A-allele among users of this drug combination ($p=1.9 \times 10^{-4}$).

4.1.2.4 Amlodipine-Atorvastatin-rs3735451 (CYP3A4)

There was an interesting association between the rs3735451 (T>C) polymorphism in *CYP3A4* gene and the concomitant use of amlodipine and atorvastatin ($p=1.3 \times 10^{-4}$). There was a 21% (1%-34%) higher chance of being a C-allele carrier among the combination users compared to non-users.

4.1.2.5 Simvastatin-Metformin- rs215095 (ABCC1)

The frequency of the G-allele at rs215095 (G(*minor allele*)>A) *ABCC1* variant was 10% (5%-16%) lower within those on the simvastatin-metformin combination ($p=2.606 \times 10^{-4}$) compared to non-users of these drugs.

4.2 Potential important associations which didn't pass Bonferroni significant level from the interacting combinations.

Under this category of results, I include associations, which although not significant after Bonferroni correction, are still of interest as they fulfil 5 criteria: Firstly, the p-value significance level was modest ($p > 0.00030$ and $<= 0.009$). Secondly, there is a predicted interaction between the two drugs (from our list of top 25 interacting combinations). Thirdly, at least one of the two drugs is a substrate for the gene relevant to the detected variant (according to DrugBank or external references). Fourthly, they occur with a genetic variant that has been previously associated with at least one drug response phenotype (according to PharmGKB). Finally, there was also a clear route of interaction consistent with the observed direction of effect.

There were 8 drug-drug-variant associations in this category from the combined Scottish cohort and other 10 results from the UKBB cohort.

Table 10, at the end of the results section, summarizes all findings under this category.

4 .3 Potential important associations which didn't pass Bonferroni significance level from the combinations with unknown routes of interactions

Under this category, I include results fulfilling the same criteria as in before, but it differs in that the drug combinations selected here belong to the top 25 combinations with no prior known mechanism to support an interaction. As explained earlier, a 'combination with unknown interaction routes' doesn't necessarily imply a completely safe combination as there might be some undiscovered routes of interactions between the two drugs. In this category, we focus on results with a moderate p-value significance level, one of the two drugs is a known substrate for the associated gene, and existence of at least one study linking the variant with any drug response phenotype. A total of 11 findings from the combined cohort and other 10 findings from the UKBB cohort have been identified under this category.

Table 10 below, summarizes all findings under this category and all other categories.

		1) Results significant after Bonferroni correction							
No.	**Drug_Combination**	**Cohort**	**Phenotype**	**SNP_id**	**Gene**	**OR**	**Lower**	**Upper**	**P.value**
1	Atorvastatin-Bisoprolol	Combined	Drug Stop	rs9516519	*ABCC4*	4.72	2.44	9.13	1.48E-05
2	Simvastatin-Metformin	Combined	Drug Stop	rs622342	*SLC22A1*	3.1	1.71	5.62	9.30E-05
3	Metformin-Gliclazide	Combined	Dose Decrease	rs1967120	*ABCC1*	1.21	1.1	1.33	9.30E-05
4	Bendroflumethiazide-Metformin	UKBB	NA	rs2199939	*ABCG2*	0.81	0.73	0.9	5.29E-05
5	Simvastatin-Furosemide	UKBB	NA	rs4148739	*ABCB1*	1.26	1.13	1.41	0.0000607
6	Atorvastatin-Metformin	UKBB	NA	rs10937158	*ABCC5*	1.13	1.06	1.2	0.000104
7	Amlodipine-Atorvastatin	UKBB	NA	rs3735451	*CYP3A4*	1.21	1.1	1.34	0.00013
8	Atorvastatin-Metformin	UKBB	NA	rs2293001	*ABCC5*	0.89	0.83	0.94	0.00016
9	Atorvastatin-Metformin	UKBB	NA	rs17731538	*ABCG2*	0.85	0.78	0.93	0.00019
10	Simvastatin-Metformin	UKBB	NA	rs215095	*ABCC1*	0.9	0.84	0.95	0.0002606
		2) Potential important associations form the **interacting** drug combinations							
No.	**Drug_Combination**	**Cohort**	**Phenotype**	**SNP_id**	**Gene**	**OR**	**Lower**	**Upper**	**P.value**
1	Simvastatin-Warfarin	Combined	Drug Stop	rs4148386	*ABCC2*	16.07	2.08	124.4	0.0005896
2	Simvastatin-Warfarin	Combined	Drug Stop	rs3814637	*CYP2C19*	9.04	2.66	30.67	0.002732
3	Atorvastatin-Bisoprolol	Combined	Drug Stop	rs4149032	*SLCO1B1*	2.95	1.51	5.75	0.001353
4	Atorvastatin-Bisoprolol	Combined	Dose Decrease	rs9561778	*ABCC4*	1.28	1.09	1.52	0.003919
5	Omeprazole-Atorvastatin	Combined	Dose Decrease	rs3814637	*CYP2C19*	1.5	1.14	1.97	0.005291
6	Amlodipine-Lansoprazole	Combined	Drug Stop	rs3814637	*CYP2C19*	2.61	1.37	4.97	0.005736
7	Atorvastatin-Lansoprazole	Combined	Dose Decrease	rs4148739	*ABCB1*	1.34	1.08	1.67	0.00851
8	Atorvastatin-Clopidogrel	Combined	Dose Decrease	rs4149118	*SLCO1B3*	1.38	1.09	1.76	0.00905
9	Simvastatin-Lansoprazole	UKBB	NA	rs4148739	*ABCB1*	1.11	1.04	1.18	0.00327
10	Simvastatin-Omeprazole	UKBB	NA	rs1128503	*ABCB1*	1.05	1.01	1.1	0.00822
11	Simvastatin-Furosemide	UKBB	NA	rs8187707	*ABCC2*	1.28	1.09	1.5	0.00398
12	Simvastatin-Furosemide	UKBB	NA	rs1045642	*ABCB1*	1.13	1.04	1.22	0.00415
13	Simvastatin-Bisoprolol	UKBB	NA	rs4149057	*SLCO1B1*	0.92	0.87	0.98	0.004773
14	Simvastatin-Amlodipine	UKBB	NA	rs3842	*ABCB1*	0.93	0.88	0.98	0.0067
15	Simvastatin-Amlodipine	UKBB	NA	rs1045642	*ABCB1*	0.95	0.92	0.99	0.00935
16	Atorvastatin-Bisoprolol	UKBB	NA	rs2306283	*SLCO1B1*	1.11	1.03	1.2	0.00853
17	Atorvastatin-Bisoprolol	UKBB	NA	rs899494	*ABCC4*	0.83	0.74	0.94	0.00238
18	Atorvastatin-Bisoprolol	UKBB	NA	rs2712807	*SLCO2B1*	0.88	0.79	0.97	0.00966
		3) Potential important associations form combinations with unknown routes of interactions							
No.	**Drug_Combination**	**Cohort**	**Phenotype**	**SNP_id**	**Gene**	**OR**	**Lower**	**Upper**	**P.value**
1	Ramipril-Metformin	Combined	Drug Stop	rs12208357	*SLC22A1*	3.13	1.5	9.32	0.0007356
2	Ramipril-Amlodipine	Combined	Dose Decrease	rs2054675	*CYP2B6*	0.61	0.44	0.84	0.0017
3	Simvastatin-Ramipril	Combined	Dose Decrease	rs34671512	*SLCO1B1*	1.82	1.3	2.55	0.00102
4	Ramipril-Bisoprolol	Combined	Drug Stop	rs3735451	*CYP3A4*	0.67	0.24	1.86	0.00127
5	Omeprazole-Atenolol	Combined	Dose Decrease	rs12248560	*CYP2C19*	1.3	1.1	1.54	0.00269
6	Ramipril-Atorvastatin	Combined	Drug Stop	rs4149032	*SLCO1B1*	2.33	1.34	4.05	0.00292
7	Atorvastatin-Metformin	Combined	Drug Stop	rs10937158	*ABCC5*	2.53	1.29	4.97	0.00462
8	Simvastatin-Perindopril	Combined	Dose Decrease	rs3814637	*CYP2C19*	1.9	1.24	2.93	0.006436
9	Ramipril-Bisoprolol	Combined	Drug Stop	rs10306135	*PTGS1*	3.25	1.47	7.2	0.0076
10	Simvastatin-Perindopril	Combined	Drug Stop	rs4148386	*ABCC2*	2.32	1.22	4.44	0.00814
11	Simvastatin-Atenolol	Combined	Drug Stop	rs622342	*SLC22A1*	0.51	0.3	0.87	0.009777
12	Simvastatin-Ramipril	UKBB	NA	rs4149118	*SLCO1B3*	1.06	1.03	1.1	0.0004548
13	Simvastatin-Metformin	UKBB	NA	rs11045879	*SLCO1B1*	0.93	0.89	0.98	0.00657
14	Simvastatin-Atenolol	UKBB	NA	rs1080985	*CYP2D6*	1.06	1.02	1.1	0.00367
15	Omeprazole-Bendroflumethiazide	UKBB	NA	rs4728709	*ABCB1*	1.17	1.04	1.31	0.00879
16	Omeprazole-Atenolol	UKBB	NA	rs2472297	*CYP1A1*	1.11	1.03	1.19	0.00669
17	Ramipril-Atorvastatin	UKBB	NA	rs4148386	*ABCC2*	1.08	1.03	1.14	0.00325
18	Bendroflumethiazide-Atorvastatin	UKBB	NA	rs3743527	*ABCC1*	0.88	0.82	0.96	0.00152
19	Omeprazole-Atenolol	UKBB	NA	rs628031	*SLC22A1*	0.91	0.85	0.98	0.00839
20	Amlodipine-Lisinopril	UKBB	NA	rs1045642	*ABCB1*	0.89	0.83	0.95	0.001134
21	Amlodipine-Lisinopril	UKBB	NA	rs12720066	*ABCB1*	1.2	1.05	1.38	0.00908

Table 10: Summary of 49 most important drug-drug-gene associations as classified into 3 categories.

5. Discussion

To the best of our knowledge, this is the first pharmacogenomic study covering the drug-drug-gene interaction topic for a large variety of commonly used drug combinations in the UK. A total of 16,200 (combined cohort) and 8,100 (UKBB cohort) results have been generated and are available to view from an online database. In the sections below, I will only focus on discussing the first two categories of associations I have presented in the results section.

5.1 Significant associations after Bonferroni correction

5.1.1 The combined cohort results (n = 3).

The 3 significant associations under this section are investigated further in chapter V, where the UK Biobank primary care data is used for replication; a more detailed discussion will also follow in chapter V for these findings.

5.1.2 The UKBB (cross-sectional) cohort results (n = 7)

5.1.2.1 Bendroflumethiazide-Metformin-rs2199939 (ABCG2)

Due to the lack of pharmacokinetic information for bendroflumethiazide, it is challenging to predict any potential pharmacokinetic route of interaction for bendroflumethiazide-containing combinations. However, it has been reported that high dose bendroflumethiazide significantly increases endogenous glucose production [192] suggesting increased potential for reducing metformin glucose-lowering effect. The minor allele (T) at rs2199939 (C>T) variant in the ABCG2 transporter is significantly depleted in patients treated with bendroflumethiazide and metformin, suggesting decreased tolerability or reduced efficacy with this drug combination. We also saw the same variant had a similar, but weaker, association for metformin-only

users from the UKBB cohort (OR = 0.97, p=0.098) in DGI study. This SNP was previously linked with increased rosuvastatin efficacy [193], suggesting reduced ABCG2 function resulted in increased rosuvastatin hepatic concertation and efficacy. However, it is unclear how ABCG2 transporter could influence bendroflumethaizde-metformin interaction.

5.1.2.2 Simvastatin-Furosemide-rs4148739 (ABCB1)

Simvastatin is a substrate for both ABCB1 and ABCC2 transporters [194,195] while furosemide is an ABCC2 inhibitor [196]. Inhibiting hepatic efflux transporters of simvastatin could results in increased efficacy, due to increased hepatic simvastatin accumulation, and increased toxicity as a result of increased plasma concentration. Both of these two phenotypes have been observed. For example, the rs717620 (C>T) SNP in the ABCC2 transporter has been reported to be associated with switching of and decreasing the dose of simvastatin therapy [197]. However, the minor reduced activity allele (A) at rs2032582 (A>C) *ABCB1* SNP has been repeatedly seen to be correlated with improved efficacy of simvastatin [198]. In the same context, another reduced function variant in *ABCB1* gene, rs4148739 (T>C), has been associated with increased antidepressant efficacy (due to decreased efflux of these drugs from the brain where these agents produce their efficacy) [199]. Consistent with this prior literature, our finding shows that simvastatin users carrying this variant allele in the ABCB1 transporter and co-treated with the ABCC2 inhibitor furosemide were more likely to be on this drug combination than people with the wild-type allele; probably due to increased simvastatin efficacy. This variant allele was only significant with users of simvastatin-furosemide combination but not with any of the two drugs individually according to our DGI study.

5.1.2.3 Atorvastatin-Metformin-rs10937158/rs2293001 (ABCC5) / rs17731538 (ABCG2) and Simvastatin-Metformin-rs215095 (ABCC1)

To date, no route of pharmacokinetic interaction has been reported between atorvastatin and metformin. However, our results suggest that such a novel interaction could exist between the two drugs. I have shown that two independent variants in the ABCC5 transporter, rs10937158 (T(*minor allele*)>C) and rs2293001(C>T) seem to be associated with the usage of this drug combination. We show that harbouring the minor allele (T) is associated with increased (rs10937158) and decreased (rs2293001) tolerability for this drug combination. A previous study has demonstrated that atorvastatin is transported by ABCC5 and that this transporter is expressed on the skeletal muscles protecting the muscles from the intracellular accumulation of statins [200]. Interestingly, in our study the rs2293001(C>T) *ABCC5* variant was associated with decreased tolerability to each drug individually (atorvastatin: OR = 0.98, p =0.03, metformin: OR = 0.96, p=0.0031), as well as in combination (atorvastatin-metformin: OR = 0.89, p=0.00016). Regarding the rs10937158 (T(*minor allele*)>C) *ABCC5* variant, this variant is also present at slightly higher frequency among metformin-only users (OR = 1.05 , p =1.9×10^{-4}), and further enriched with when combined with atorvastatin (OR = 1.13 , p=1.0×10^{-4}). In keeping with the direction of effect we observed (increased tolerability), the T allele at this variant has also been reported to be linked to decreased irinotecan-induced severe diarrhoea [201].

Furthermore, genetic variants in ABCG2 transporter markedly alter atorvastatin plasma concentration [202]. In this DDGI study, the A-allele at rs17731538 (G>A) in *ABCG2* gene is at a lower frequency in atorvastatin-metformin users than non-users suggesting decreased tolerability or efficacy. Consistent with this, the A-allele at this

variant has also been linked with decreased methotrexate therapeutic efficacy [203]. Although we are not aware of any pharmacokinetic route of interaction between the two drugs, our results outlined here merit further investigation in in-vitro systems.

In addition, as mentioned earlier, there is no known route of interaction between simvastatin and metformin. However, our DGI study reveals that the minor allele (G) at rs215095 (G(*minor allele*)>A) variant in the ABCC1 transporter was associated with reduced tolerability to each drug individually (simvastatin : OR = 0.98 , p=0.071 , metformin : OR = 0.94 , p=0.0034) before observing a larger effect when both drugs are combined together (simvastatin-metformin : OR = 0.90 , $p=2 \times 10^{-4}$). Interestingly, this is the second metformin-containing combination showing a correlation with the ABCC1 transporter in an independent variant and toward the same direction of effect (increased likelihood to decrease the dose).

5.1.2.4 Amlodipine-Atorvastatin-rs3735451 (CYP3A4)

Both atorvastatin and amlodipine are extensively metabolized in the liver with the CYP3A4 being the major metabolic pathway of both agents [204,205]. This suggests the potential for a competitive inhibition of the victim drug (atorvastatin) by the perpetrator drug (amlodipine) due to a shared route of elimination. In a drug-drug interaction study presented by Food and Drug Administration (FDA), atorvastatin did not change the pharmacokinetics of amlodipine while the latter resulted in an 18% increased atorvastatin plasma drug concentration which was deemed not to be clinically significant [206]. However, the finding from our study suggests that genetic variation in CYP3A4 could affect the therapeutic efficacy of atorvastatin when combined with amlodipine. The C-allele frequency of the

rs3735451 (T>C) *CYP3A4* variant in patients on this drug combination were higher compared to those not on the combination suggesting a beneficial drug-drug-gene interaction. A couple of pharmacogenomic studies have linked *CYP3A4* polymorphisms with atorvastatin efficacy and pharmacokinetics. Carriers of the reduced activity allele T of rs35599367 (C>T) in *CYP3A4* required lower atorvastatin dose compared to non-carriers [207]. In contrast, the *CYP3A4* gain-of-function variant rs2740574 (C>T) was associated with a low risk to reduce atorvastatin dose, low risk to switch it to an alternative drug, or increased atorvastatin plasma concentration [208]. Here, we also report another *CYP3A4* variant (rs3735451 (T>C)) which was significant with the use of both drugs together but not atorvastatin alone. This is predicted to be a reduced activity variant as a previous report shows it is correlated with increased methadone (a CYP3A4 substrate)-related side effects [209].The exact mechanism of the atorvastatin-amlodipine-rs3735451 (T>C) interaction remains uncertain, but it seems likely that the combined competitive inhibitory effect of amlodipine on CYP3A4 and the reduced CYP3A4 activity (by rs3735451 (T>C) variant) would result in elevated hepatic atorvastatin concentration leading to the increased therapeutic efficacy of atorvastatin since the liver is its site of action.

5.2 Potential important associations which didn't pass Bonferroni significance level from the interacting combinations

5.2.1 The combined cohort results (n = 8).

5.2.1.1 Simvastatin-Warfarin-rs4148386 (ABCC2) /rs3814637(CYP2C19)

The interaction route between simvastatin and warfarin is complex as both drugs share multiple elimination pathways and can inhibit each other in multiple pathways. Both drugs are substrates for CYP3A4, CYP2C8, and CYP2C19 with warfarin being an inhibitor for CYP2C8 and simvastatin being a substrate for ABCB1 and ABCC2 transporters and an inhibitor for CYP2C8 and CYP2C9 enzymes [210-212]. These complex interaction routes suggest that toxicity could be seen from either of the two drugs. For example, a previous report has shown that warfarin was one of the common interacting drugs among patients who developed statin-induced rhabdomyolysis and that simvastatin was the most frequently reported statin among the affected individuals [213]. Genetic variability could influence this risk. I have shown that minor allele (G) at rs4148386 (G (*minor allele*)>A) in ABCC2 was associated with increased risk for stopping either of the two drugs combined use. This could result from the double inhibitory effects on two simvastatin elimination pathways: ABCC2 (by the G allele which was reported previously to be linked with decreased carbamazepine clearance [214] suggesting reduced ABCC2 function) and CYP2C8 (by warfarin treatment) resulting in increased simvastatin toxicity.

On the other hand, it has been illustrated that warfarin users co-treated with simvastatin have a moderate increase in warfarin-induced bleeding [215]. It has been reported that carriers of the *CYP2C9**3 reduced function variant required lower warfarin doses compared to non-carriers when co-treated with simvastatin [216]. Similarly, our results identified that a variant, rs3814637 (C>T), in another warfarin

elimination pathway (CYP2C19) is linked with stopping any of the two drugs during their combined use. It is worth noting that this variant was previously linked with the need for lower warfarin doses [217] which is consistent with our finding. The difference, however, is that in the DDGI study this variant is linked with the drug-stop phenotype only where there is concomitant use of the two drugs but not with warfarin-only users (see DGI study results).

5.2.1.2 Atorvastatin-Bisoprolol-rs4149032 (*SLCO1B1*) / rs9561778 (*ABCC4*)

Atorvastatin and bisoprolol could interact through the CYP3A4 enzyme pathway. Bisoprolol is metabolized by the CYP3A4 enzyme [218] while atorvastatin is an inhibitor for CYP3A4 [219] suggesting potential changes in the pharmacokinetics of bisoprolol due to both atorvastatin use and the *CYP3A4* genotype. We identified two genetic variants, rs4149032 (C>T) and rs9561778 (G>T), in SLCO1B1 and ABCC4, respectively that seemed to influence this interaction. The former SNP has been previously reported to decrease letermovir and rifampicin (SLCO1B1 substrates) systemic exposure [220,221]; two findings suggesting the variant is gain-of-function for SLCO1B1 hepatic uptake activity. I found that this SNP is linked with increased odds of stopping one of the two drugs, bisoprolol or atorvastatin, during their combined use. I propose that increased atorvastatin hepatic uptake by SLCO1B1 will result in increased atorvastatin hepatic accumulation, which in turn, could result in an increased CYP3A4 inhibitory effect and therefore, increased bisoprolol toxicity even though this could be also associated with increased efficacy. The other *ABCC4* variant (rs9561778) has been reported to increase cyclophosphamide side effects [222]. Cyclophosphamide is a prodrug which is heavily activated and metabolized in the liver [222]. This drug and its active metabolites are transported by ABCC4 [222]. ABCC4 is expressed in the basolateral hepatic membrane facilitating the efflux of

xenobiotics into the systemic circulation. One proposed mechanism for the increased cyclophosphamide toxicity is that the rs9561778 (G>T) variant might be associated with increased ABCC4 function and increased systemic cyclophosphamide exposure accordingly. We show that carriers of this variant who were on the atorvastatin-bisoprolol combination were at increased risk of a decrease in the daily dose of either of the two drugs when taken concomitantly. One possible mechanism is that the combined effect between CYP3A4 inhibition by atorvastatin and the SNP associated increased efflux by ABCC4 could result in increased bisoprolol plasma concentration and toxicity. None of the two variants significantly altered drug response phenotype with each of the drugs individually, only when combined.

5.2.1.3 Omeprazole-Atorvastatin-rs3814637 (CYP2C19) , Lansoprazole-Amlodipine-rs3814637 (CYP2C19), and Lansoprazole-Atorvastatin-rs4148739 (ABCB1)

Omeprazole and lansoprazole are primarily metabolized by CYP3A4 and CYP2C19 enzymes with some contribution of CYP2C9 and CYP2C8 [223,224]. The two drugs are also substrates and inhibitors for ABCB1 transporter [223,224]. Atorvastatin is a substrate and inhibitor for CYP3A4 and ABCB1; it is also an inhibitor for CYP2C8, CYP2C9 and CYP2C19 [219,225] while amlodipine is a substrate and weak inhibitor for CYP3A4 enzyme [205,226]. The reduced activity variants *CYP2C19*2/*3* have been repeatedly correlated with increased omeprazole/lansoprazole therapeutic efficacy [227,228]. *CYP2C19*3* is extremely rare in the European ancestry, but even for the common *CYP2C19*2* variant, we couldn't see this SNP to be significantly correlated with drug-stop or dose-decrease phenotype for any of the two drugs. Nevertheless, our findings show another independent low activity *CYP2C19* SNP (rs3814637 (C>T)) seemed to be associated with increased toxicity from these two proton pump inhibitors (PPIs) when they are co-administered with other interacting

drugs. Carriers of this variant who were either treated with omeprazole-atorvastatin or lansoprazole-amlodipine combinations were more likely to decrease the dose of or stop any of the two drugs during their combined use respectively. The existence of other risk factors such as atorvastatin (a CYP3A4/CYP2C19 inhibitor) or amlodipine (a CYP3A4 inhibitor/competitive inhibitor) could have resulted in increased PPIs-related side effects.

As outlined above, lansoprazole and atorvastatin are both substrates and inhibitors for ABCB1 transporters suggesting that the victim drug could be either of the two drugs. As with the other two PPI-containing combinations above, no pharmacokinetic drug-drug interaction studies have been reported for lansoprazole and atorvastatin. However, our results demonstrate that those with the low activity ABCB1 variant rs4148739 (T>C) had higher odds to reduce the dose of either lansoprazole or atorvastatin when the two drugs are used concomitantly.

5.2.1.4 Atorvastatin-Clopidogrel-rs4149118 (SLCO1B3)

The interaction between atorvastatin and clopidogrel could be primarily mediated via CYP3A4 competitive inhibition pathway as both drugs are substrates for CYP3A4 [204,155]. One previous study reported that the percentage of cardiovascular events was higher among users of both atorvastatin and clopidogrel compared to users of clopidogrel alone [229]. This clinical outcome could occur as a result of reduced efficacy of either of the two drugs. The authors have suggested that the mechanism could be as a result of the competitive inhibition of CYP3A4 by atorvastatin resulting in decreased activation of clopidogrel into its active metabolites. However, an alternative explanation could be mediated via SLCO1B3. I have shown that carriers of the minor allele (G) at rs4149118 (G>A) in SLCO1B3 were more likely to decrease the dose of atorvastatin or clopidogrel during their combined use. In a previous study,

the G-allele has been reported to be associated with decreased clearance and increased plasma concentration of docetaxel (a SLCO1B3 substrate) indicating reduced hepatic uptake of the drug [230]. It could be the case that decreased atorvastatin uptake into the liver in addition to the competitive CYP3A4 inhibition by clopidogrel have resulted in increased atorvastatin plasma concentration and reduced efficacy. Increased toxicity could lead to dose reduction (according to our finding), and reduced efficacy could lead to increased cardiovascular events as has been shown in the study above.

5.2.2 The UKBB (cross-sectional) cohort results (n=10).

All results under this category represent statin-containing combinations. There are 7 simvastatin- and 3 atorvastatin-containing associations which are discussed below.

5.2.2.1 Simvastatin-Lansoprazole-rs4148739 (ABCB1) and Simvastatin-Omeprazole-rs1128503 (ABCB1)

As mentioned previously, simvastatin is a substrate for CYP3A4, CYP2C8, and CYP2C19 enzymes and is transported by ABCB1 and ABCC2 transporters. The drug is an inhibitor for CYP2C8, CYP2C9, and ABCB1. As also previously mentioned, lansoprazole and omeprazole are metabolized by CYP3A4 and CYP2C19 enzymes and are substrates for ABCB1 transporter. The two drugs are also inhibitors for *CYP3A4* and *ABCB1* genes. The elimination pathways presented here for simvastatin and these two PPIs suggest that they can interact with each other via multiple routes either by direct or competitive inhibition, mainly mediated by CYP3A4, CYP2C19, and ABCB1 pathways. In one previous study [231], PPIs use was shown to have a modest increase of statin lipid-lowering efficacy. In keeping with this, the results from

our DDGI study shows that carriers of the *ABCB1* variants: rs4148739 (T>C) or rs1128503 (A(*minor allele*)>G) are more likely overrepresented in the simvastatin-lansoprazole or simvastatin-omeprazole combinations respectively. In fact, in two previous studies [232,233], the latter variant allele (A allele at rs1128503 SNP), has been associated with increased low-density lipoprotein (LDL) reduction among simvastatin users. However, the results from our DGI study from the UKBB found no association between this variant and efficacy/tolerability of simvastatin (OR = 1.01, p=0.41). The association was only significant among simvastatin users who are co-treated with omeprazole. Thus, the evidence indicates the existence of a favourable interaction between simvastatin and PPIs which may be consequent to reduced ABCB1 function by PPIs and/or carriage reduced activity variants in ABCB1 transporter, in addition to the known CYP3A4/CYP2C19 competitive inhibition by PPIs on simvastatin elimination pathway. Statin-PPI interactions need to be further investigated considering different clinical outcomes such as the effect of statins on PPIs safety and efficacy.

5.2.2.2 Simvastatin-Furosemide-rs8187707 (ABCC2)/ rs1045642 (ABCB1)

As outlined previously, furosemide could interact with simvastatin by inhibiting the hepatic efflux transporter ABCC2 which contributes to simvastatin biliary excretion. I have also reported a DGI study linking the rs717620 (T>C) *ABCC2* SNP with increased odds of simvastatin dose reduction, although in UK Biobank we did not see such an association. However, in the DDGI study, I have identified another independent *ABCC2* variant which is associated with the opposite direction of effect when simvastatin is co-administered with furosemide. Carriers of the rs8187707 (C>T) variant in ABCC2 transporter seemed more tolerant to this drug combination as indicated by enrichment of this variant in this treatment group. This SNP has been

previously been reported to be associated with increased tenofovir renal-related adverse reactions [234]. It has also been reported that tenofovir is a substrate for the ABCC2 transporter and that inhibition of this transporter by ritonavir increased tenofovir-related toxicity [235]. These observations could indicate that the rs8187707 (C>T) is a reduced activity variant. As a result of this, this reduced function variant could combine with furosemide to inhibit ABCC2, resulting in reduced simvastatin hepatic efflux and, in turn, increased efficacy which could explain our observation. There could be a similar explanation for those carrying the rs1045642 (A>G) *ABCB1* variant and co-treated with the ABCC2 inhibitor furosemide.

5.2.2.3 Simvastatin-Bisoprolol-rs4149057 (SLCO1B1)

As discussed previously, bisoprolol is mainly metabolized by CYP3A4; which implies that simvastatin and bisoprolol could interact with each other via competitive inhibition on CYP3A4 enzyme. The *SLCO1B1* variant, rs4149057 (T(*minor allele*)>C, *SLCO1B1*18*), has been reported to be associated with decreased rosuvastatin (a SLCO1B1 substrate) hepatic uptake [236]. Our results show that this variant was found in a lower frequency within individuals treated with simvastatin and bisoprolol. Both reduced simvastatin hepatic uptake and the competitive inhibitory effect of bisoprolol on CYP3A4 could result in increased simvastatin plasma levels and toxicity as well as a loss of its efficacy. These phenotypes could explain our finding that the *SLCO1B1*18* variant is linked to decreased tolerability to this drug combination.

5.2.2.4 Simvastatin-Amlodipine-rs3842/rs1045642 (ABCB1)

Both simvastatin and amlodipine are metabolized by CYP3A4, suggesting a competitive inhibition could result between the two drugs. Results from drug-drug interaction studies demonstrate that amlodipine has significantly increased the plasma

concentrations of simvastatin [237]. Therefore, the Medicines and Healthcare Products Regulatory Agency (MHRA) has recommended reducing simvastatin dose from 40 mg to 20 mg for patients who are co-treated with amlodipine [238] to avoid simvastatin-induced myopathies. Amlodipine can also interact with simvastatin via the ABCB1 transporter as it is a substrate and inhibitor for this transporter [239,240]. In the UK Biobank DDGI study, the minor alleles of the *ABCB1* variants rs3842 (T>C) and rs1045642 (A>G) were depleted among users of this drug combination suggesting decreased tolerance and/or efficacy. rs3842 has been reported to be associated with increased systemic exposure to efavirenz, suggesting decreased ABCB1 activity [241] while rs1045642 has shown conflicting results regarding its functional impact on simvastatin pharmacokinetics [242]. It might be the case that the existence of two risk factors (CYP3A4 inhibition by amlodipine + ABCB1 reduced activity) have resulted in increased simvastatin toxicity leading to the decreased tolerability we have observed. It is worth noting that none of these two variants was significant with simvastatin-only users from the UKBB (OR = 0.98, p= 0.1096, and OR = 0.99, p=0.1408 respectively) and were only significant with those on simvastatin and amlodipine combination.

5.2.2.5 Atorvastatin-Bisoprolol-2306283 (SLCO1B1)/ rs2712807 (SLCO2B1)/ rs899494 (ABCC4)

As outlined previously, the two drugs could interact via CYP3A4 route. Interestingly, our results from both the combined and the UKBB cohorts show multiple lines of evidence that the use of atorvastatin-bisoprolol combination could be influenced by genetic variability in the ABCC4 transporter and SLCO family of transporters. Firstly, for the ABCC4 transporter from the combined cohort, I have shown a significant

association after Bonferroni correction between the rs9516519 variant in ABCC4 transporter and stopping any of the two drugs when used concomitantly. Secondly, another independent *ABCC4* variant (rs9561778) from the same cohort was also observed to be associated with increased risk of reducing the dose of any of the two drugs during their interaction. Thirdly, I have identified another independent *ABCC4* SNP (rs899494 (A(*minor allele*)>G)) from an independent cohort (UKBB) to be linked with decreased tolerability to this combination.

For the SLCO family, I have shown, firstly, an association of rs4149032 *SLCO1B1* SNP and stopping any of the two drugs when used together from the combined cohort. Secondly, from the UK Biobank, I showed that the minor allele (G) at rs2712807 (G(*minor allele*)>A) of SLCO2B1 was underrepresented in this treatment group suggesting increased side effects or reduced efficacy. Thirdly, I have also shown in the UK Biobank that carriers of the rs2306283 (A>G) *SLCO1B1* variant seemed more tolerant to this combination. This variant has been previously reported to be associated with increased atorvastatin lipid-lowering efficacy [243]. However, another study found no association with atorvastatin response [244]. Our DGI study also shows no association between this variant and tolerability to atorvastatin.

6. Conclusion

In the light of the very limited knowledge we have to date on drug-drug-gene interactions (only 33 publications in the field according to our review in chapter I), the present work has provided for the first time a large-scale coverage of clinical pharmacogenomic associations between 162 genetic variants in important enzymes and transporters and 50 commonly used chronic drug combinations using 3 different drug response phenotypes (drug stop, dose decrease, and changes in genotype distribution among drugs' users) in two cohorts (the Scottish cohort + the UKBB cohort). I have generated a total of 24,300 drug-drug-genetic variants associations results which are accessible online via https://c1abo933.caspio.com/dp/d81f7000033513b3ab1c4431b5a2 .

I have discovered a total of 10 novel drug-drug-genetic variant associations. I have also identified 18 potentially important novel associations within the list of interacting drug combinations. Finally, I also show other 21 potentially interesting associations between drugs not known to interact according to current knowledge. These might suggest novel routes of interactions which will require further investigation in the future.

However, replication of DDGI results will be more challenging than that of DGI findings due to the small sample sizes in general requiring very large cohorts and the complexity of the phenotypes themselves. Instead of replication, validation of the findings by in-vitro studies would provide supportive evidence that our clinical findings are real and potentially clinically actionable.

Chapter V:

Replication Results from the UKBB Primary Care Data for the top Drug-Gene and Drug-Drug-Gene Interactions

Abstract

In this chapter, I undertake replication of the top 9 DGI results (8 passed Bonferroni significance level and 1 being the closest association to the Bonferroni significance level occurring in an interesting gene) and 3 DDG interactions findings from the combined cohort using the UKBB primary care data.

Out of 9 novel DGIs associations; 3 associations have been replicated, validated, or have supporting evidence. **1)**The C allele at rs4918758 (T>C) *CYP2C9* SNP was associated with 30% (15%-40%; p= 8×10^{-5}) and 19% (1%-33%; p=0.037) lower odds of quinine daily dose reduction for the discovery and the replication cohorts respectively (replicated). 2) Carriage of the A allele at rs9895420 (T>A) *ABCC3* variant was associated with a 46% (24%-62%) reduced odds for doxazosin dose reduction (p= 1.2 × 10-4, discovery) and 1.0mmHg greater reduction in systolic blood pressure (p =0.0089, validation). 3) In addition, the *CYP2D6*2* variant representing an extensive metabolizer phenotype was observed to be linked with a 30% (18 %-40%) reduction in odds of stopping ramipril treatment (p = 1.01 × 10-5, discovery) while the *CYP2D6*4* variant representing a poor metabolizer phenotype was associated with a 29% (8%-54%) increased odds of stopping the drug (p= 0.00654, supporting evidence).

We have also detected one finding to be directionally consistent with a p-value close to the significance level in the replication cohort. The C allele at rs868853 *ABCC4* variant was correlated with reduced odds for decreasing amlodipine daily dose in the discovery (OR = 0.55, p = 2.8×10^{-4}) and the replication cohort (OR = 0.86, p = 0.092). We have also noted one other finding (Clopidogrel-rs12353214 (C>T,

PTGS1(COX-1))-drug-stop) to be close to the Bonferroni significance level in the discovery cohort (OR = 0.57; p= 5.3×10^{-4}); this was also close to significance level in the replication cohort (OR = 0.86; p= 0.067).

In addition, 3 novel DDGIs were identified in the discovery analysis with one association supported (validated) using an alternative phenotype in UK Biobank but not replicated with the same phenotype. In the discovery cohort, the G allele at rs9516519 (T>G) variant in the ABCC4 transporter was associated with 4.72 (2.44-9.13) times increased risk to stop bisoprolol or atorvastatin treatments when they are used concomitantly (p=1.48×10^{-5}). Although this finding did not replicate directly, in the replication cohort, the drug combination alone was associated with great SBP reduction (~ 8 mmHg drop in mean SBP ($p < 2 \times 10^{-16}$)). When the genotype data for rs9516519 (T>G) SNP was included, it seems that it has an added effect on SBP reduction. Compared to the normal genotype (TT), the heterozygous (TG) and recessive (GG) genotypes have shown a clear trend toward increased SBP drop among the combination users with the difference being - 2.46 and - 4.54 mmHg for both genotypes respectively (p (trend) = 0.02242).

1. Introduction

There have been numerous pharmacogenomic associations which are published without replication. It should also be noted that even for the studies where there have been attempts to replicate findings, only few associations have been successfully replicated. In fact, in one large study assessing the replication rate in genetic disease studies for publications from 2007 until 2010, authors have found that only 13 out of the 1151 (~ 1%) genotype-phenotype associations were successfully replicated [245]. This led to a clear move in disease genetics for a requirement for robust replication for all published studies. This has yet to fully translate into pharmacogenetics, however, as the findings outlined in chapters III and IV use a generic response phenotype across many drugs, there is clearly a need to establish which of the results replicates in an independent cohort. This was not possible for us until the release of the UK Biobank primary care data late in 2019. Therefore, in this chapter, we will attempt to replicate the top 9 drug-gene and 3 drug-drug-gene associations which were initially discovered in the combined Scottish cohort by utilizing an independent larger cohort, the UKBB primary care data.

2. Methodology

I used the UKBB primary care data in order to investigate the replicability of our top DG and DDG interactions findings from the combined Scottish cohort. Primarily, the same phenotype was used (i.e. drug-stop or dose-decrease, see chapters III and IV for the definition of these phenotypes) to check for replication. However, for some findings, I undertook a further investigation to examine a specific hypothesis related to the particular drug in question. A number of the drugs implicated were antihypertensives, so I used the reduction in mean systolic blood pressure due to the

DGI or DDGI. Drug response phenotype definitions for these two scenarios were explained in detail in chapter II. Here, I briefly re-describe them.

2.1 Drug-gene interaction-related systolic blood pressure changes.

After calculating the mean of the SBP measures 1-year before and 1-year after the antihypertensive agent, a multiple regression model, is then used to study the influence of the genetic variant of interest on the blood pressure. The outcome (Y) continuous variable was the mean of the SBP measures 1-year after treatment. The explanatory variables were mean SBP one-year pre-treatment (to adjust for the drug-only effect), the SNP of interest as a 3-level categorical variable, age and sex.

2.2 Drug-drug-gene interaction-related systolic blood pressure changes.

Two approaches were used to investigate drug-drug-gene interactions for systolic blood pressure (SBP) reduction. In the first approach, only those who started the antihypertensive drug first then added the perpetrator drug are included. For these patients, a mixed effect model was utilized to identify whether there was a significant difference in SBP during treatment with the antihypertensive drug alone and during treatment with the combination (antihypertensive drug + perpetrator drug) for the same patient.

In the second approach, in order to increase the sample size, two groups of patients were compared: 1) all patients starting the antihypertensive drug alone and 2) those who started the perpetrator drug first before adding the antihypertensive drug. In this approach, a multiple regression model adjusted by SBP before starting the antihypertensive agent, age, and sex is run to identify whether there was a significant difference in mean SBP levels between the two groups (i.e. during bisoprolol-only SBP vs during interaction SBP). The analysis is run first without the genotype data to

investigate whether there is a significant drug-drug interaction. Then, the genotype data is included to identify whether there is a significant drug-drug-gene interaction.

3. Results

3.1 The replication results for the top 9 drug-gene interactions findings

Table 11 at the end of DGI results section summarizes all results under this category.

We present these results in order starting from the ones with the best evidence.

3.1.1 Replicated findings (n = 1).

Quinine-rs4918758 (T>C, *CYP2C9*): low likelihood to decrease the dose.

This association was observed in GoDARTs and GoSHARE cohorts with a consistent direction of effect, with the Generation Scotland results being non-significant (p=0.835). GoDARTs subjects with this rs4918758 *CYP2C9* variant experienced a 28% (11%- 41%) lower odds of quinine dose decrease (p=0.00160). A similar scenario was also seen with GoSHARE participants carrying this variant who experienced a 50% (23%-67%) per-allele lower odds of quinine dose decrease ($p = 8.14 \times 10^{-4}$).

I then investigated the replicability of this finding in the UKBB primary care data, where a similar picture was seen. The quinine group carrying the C allele had a 19% (1%- 33%) lower odds of a decrease in quinine dose per allele (p= 0.0374, log-additive model). Compared to TT homozygotes, the heterozygous genotype (TC) was associated with a 10% lower odds of quinine dose reduction, and those CC at this variant had a 41% lower odds of dose reduction (p= 0.0660, codominant model).

After meta-analyzing both the discovery and the replication findings, carriers of the variant allele were 25% (15%-44%) less likely to reduce quinine daily dose (p=1.64 × 10-5). Interestingly, consistent with this finding, our data also show that the other two known loss-of-function variants rs1799853 (C>T, *CYP2C9*2*) and rs1057910 (A>C, *CYP2C9*3*) are associated with a 25% (4%-42%, p= 0.017) and 31% (1%-51%, p=0.0325) lower tendency to reduce the daily dose of quinine respectively.

3.1.2 Validated findings (n = 1).

Doxazosin-rs9895420 (T>A, *ABCC3*): low likelihood to decrease the dose.

The rs9895420 A-allele was associated with lower odds of doxazosin dose-decrease in GoDARTs (OR = 0.51 (0.33-0.77); p = 0.00477). A similar direction of effects is seen among GS and GoSHARE participants (29% and 43% lower odds for dose reduction) even individually, the associations in these cohorts were not significant.

In the UKBB primary care data, there was a similar trend of association as this variant showed a 10% less chance for dose reduction, but this result was not statistically significant (p=0.33857).

The above observations suggest increased tolerability in term of side effects. I then examined the effect of this variant on SBP reduction. Interestingly, there was an association between the A-allele at rs9895420 and increased doxazosin efficacy (p=0.008946). The A allele was associated with a 1.037 mmHg increase in mean SBP reduction 1 year after doxazosin. Compared to TT homozygotes, the SBP reduction was 1.03 mmHg greater in TA heterozygotes and 2.13 mmHg greater in AA homozygotes. (see Table 12 below).

Influence of rs9895420 (T>A) variant on doxazosin SBP lowering efficacy

	n	mean	St. error	difference	lower	upper	p-value
Codominant							
A/A	4326	147	0.2193	0			0.032808
A/a	1023	146.1	0.4394	-1.031	-1.873	-0.1896	
a/a	44	145	1.7791	-2.132	-5.8	1.5356	
log-Additive				-1.037	-1.814	-0.2598	0.008946

Table 12: rs9895420 (T>A, *ABCC3*)-genotype-based changes in mean SBP 1 year after doxazosin adjusted by mean SBP 1 year before doxazosin, age, and sex.

3.1.3 Findings with supporting evidence (n =1).

Ramipril-rs1135840 (G>C, *CYP2D6*): low likelihood to stop treatment.

This association was significant in all of the 3 cohorts individually with a consistent direction of effect. Per C-allele at rs1135840, ramipril users carrying this variant from GoDARTs , GS and GoSHARE cohorts were 25% (8%-39% , p = 0.005417) , 38% (9%-58%, p = 0.01287) , and 34% (9%-40%, p = 0.008746) less likely to stop ramipril treatment.

Unfortunately, the rs1135840 (*CYP2D6*2)* SNP, representing the extensive metabolizer phenotype, deviated markedly from HWE in the UKBB cohort and therefore cannot be used. However, in the combined cohort another SNP rs3892097 (*CYP2D6*4*), representing a poor metabolizer phenotype, was associated with stopping ramipril therapy (OR = 1.29 (1.08-1.54)), p=0.00654) providing more evidence on a potential ramipril-CYP2D6 correlation. Further support for this finding can be seen with other angiotensin converting enzyme (ACE) inhibitors in our data. The *CYP2D6*4* variant was also associated with a 31% (1%-72%) increased odds of decreasing enalapril daily dose (p=0.053) while the *CYP2D6*2* variant was associated with a 13% (1%-23%) lower risk to decrease lisinopril daily dose (p=0.032).

3.1.4 Findings where their replications were not significant but directionally consistent with a p-value close to the significance level (n =2).

1) Amlodipine-rs868853(C(*minor allele*)>T, *ABCC4*): decreased likelihood to decrease the dose.

There was a large and consistent effect of rs868853 on amlodipine response in all of the 3 Scottish cohorts. Amlodipine users from GoDARTs, GS, and GoSHARE carrying the C-allele were 44%, 52%, and 45% less likely to decrease amlodipine daily dose per allele. This finding was significant for GoDARTs cohort (p=0.0060) and after combining all cohorts together ($p=2.8 \times 10^{-4}$). This drug-gene correlation has then been examined using the UKBB primary care data. There was a similar association of rs868853 and likelihood to lower amlodipine dose with C-allele carriers having a 14% lower odds to reduce amlodipine daily dose, although the result was not significant in the UKBB cohort (p=0.092, log-additive; p=0.069 dominant).

2) Clopidogrel-rs12353214 (C>T, *PTGS1*): decreased likelihood to stop the drug.

In GoDARTs, GS, and GoSHARE, the T-allele was associated with being 24% (p=0.044), 58% (p=0.34), and 54% (p=0.12) less likely to stop clopidogrel treatment. This finding was found close to the Bonferroni significance level after combing all cohorts together ($p=5.3 \times 10^{-4}$) as outlined earlier.

Examining this association in a larger UKBB cohort showed a similar, although not-statistically significant effect: carrying the T-allele was associated with a 14% lower tendency to stop clopidogrel therapy (p=0.0671).

3.1.5 Other non-replicated/non-validated associations (n = 4).

1) Nifedipine-rs152023 (C(*minor allele*)>T, *ABCC1*): increased likelihood to stop the drug.

This finding was significant in the GoDARTs cohort (OR 1.37 (1.15-1.64), p = 5.184 $\times 10^{-4}$). The results from GS and GoSHARE cohorts showed consistent directions of effects even though they were not significant (OR 1.07 (0.77-1.47), p=0.6961, and OR 1.43 (0.97-2.1), p=0.06848 respectively). However, a further investigation from the UKBB primary care data for this association showed no significant association between rs152023 and stopping nifedipine treatment (p=0.9825).

2) Valproic acid-rs7916649 (T>A, *CYP2C19*): increased likelihood to decrease the dose.

The results from GoDARTs and GS cohorts show a significant association between the rs7916649 A-allele and being 2.03 (1.23-3.35) and 2.01 (1.07-3.78) times more likely to decrease valproic acid daily dose (p-values = 0.0044 and 0.027 respectively). The GoSHARE result was directionally consistent (OR=1.62 (0.74-3.54), p = 0.22). However, there was no replication of these results in the UKBB primary care data (OR = 0.95, p = 0.732).

3) Metformin-rs1045642 (A>G, *ABCB1*): low likelihood to stop the drug.

GoDARTs , GS, and GoSHARE individuals on metformin and carrying the G-allele at rs1045642 were 29% , 28%, and 17% less likely to stop metformin treatment. The only cohort individually significant was GoDARTs (p=0.000717). This result did not replicate in UK Biobank.

4) Nicorandil-rs5788 (C>A, *PTGS1*): increased likelihood to stop the drug.

In the three Scottish cohorts, there was an increased odd of stopping nicorandil in A-allele carriers at rs5788 for the GoDARTs and GoSHARE cohorts (ORs = 2.57 and 1.37 respectively); the GS result was not directionally consistent and was non-significant (OR = 0.55, p=0.35). This result did not replicate in UK Biobank.

Table 10 below summarizes the results of our top 9 associations from the combined cohort and their replication/validation results from the UKBB primary care data.

This page was left intentionally

					GoDARTs		GS		GoSHARE		Combined (discovery)		UKBB (replication/validation)	
No.	drug	gene	SNP	Phenotype	OR (CI)	p-value	OR (CI)	p-value	OR (CI)	p-value	OR (CI)	p-value	OR (CI)/difference(CI)	p-value
1	Ramipril	*CYP2D6*	rs1135840 (G>C,*CYP2D6*2*)	Drug stop	0.75 (0.61-0.92)	0.005417	0.62 (0.42-0.91)	0.01287	0.66 (0.49-0.91)	0.008746	**0.7** (0.6-0.82)	**1.01E-05**	1.04 (0.98-1.11)	0.2017
1	Ramipril	*CYP2D6*	rs3892097 (C>T,*CYP2D6*4*)	Drug stop	1.29 (1.03-1.63)	0.03301	1.76 (1.12-2.76)	0.01581	1.07 (0.74-1.55)	0.7295	1.29 (1.08-1.54)	**6.54E-03**	NA	NA
2	Quinine	*CYP2C9*	rs4918758 (T>C)	Dose decrease	0.72 (0.59-0.89)	0.001602	1.05 (0.65-1.7)	0.8385	0.5 (0.33-0.77)	0.000814	**0.7** (0.6-0.85)	**8.00E-05**	**0.81** (0.67-0.99)	**0.0374**
3	Nifedipine	*ABCC1*	rs152023 (C(*minor*)>	Drug stop	1.37 (1.15-1.64)	0.000518	1.07 (0.77-1.47)	0.6961	1.43 (0.97-2.1)	0.06848	**1.32** (1.15-1.53)	**1.17E-04**	1 (0.91-1.1)	0.9825
4	Doxazosin	*ABCC3*	rs9895420 (T>A)	Dose decrease	0.51 (0.33-0.77)	0.00477	0.71 (0.28-1.78)	0.4418	0.57 (0.26-0.24)	0.1264	**0.54** (0.38-0.76)	**1.20E-04**	**0.9** (0.73-1.12)	0.33857
4	Doxazosin	*ABCC3*	rs9895420 (T>A)	SBP changes	NA	NA	NA	NA	NA	NA	NA	NA	**-1.037** (-1.814 , -0.25)	**0.008946**
5	Valproic acid	*CYP2C19*	rs7916649 (G>A)	Dose decrease	2.03 (1.23-3.35)	0.004444	2.01 (1.07-3.78)	0.02728	1.62 (0.74-3.54)	0.2282	**1.95** (1.37-2.76)	**1.48E-04**	0.95 (0.71-1.28)	0.7324
6	Metformin	*ABCB1*	rs1045642 (A>G , ABCB1*2/13)	Drug stop	0.71 (0.58-0.87)	0.000717	0.72 (0.46-1.13)	0.1505	0.83 (0.63-1.09)	0.1715	**0.75** (0.64-0.87)	**1.64E-04**	**0.96** (0.83-1.11)	0.579
7	Nicorandil	*PTGS1*	rs5788 (C>A)	Drug stop	2.57 (1.72-3.84)	1.01E-05	0.55 (0.14-2.09)	0.3579	1.37 (0.69-2.72)	0.1729	**1.93** (1.39-2.69)	**1.72E-04**	**1.09** (0.86-1.37)	0.4779
8	Amlodipine	*ABCC4*	rs868853 (C(*minor*)>	Dose decrease	0.56 (0.39-0.81)	0.006059	0.48 (0.15-1.53)	0.5257	0.55 (0.28-1.08)	0.2551	**0.55** (0.4-0.75)	**2.80E-04**	**0.86** (0.71-1.03)	**0.09279**
9	Clopidogrel	*PTGS1*	rs12353214 (C>T)	Drug Stop	0.67 (0.44-1.01)	0.04453	0.42 (0.15-1.18)	0.3435	0.57 (0.41-0.8)	0.1283	**0.57** (0.41-0.8)	**0.00053**	**0.86** (0.72-1.01)	**0.067144**

Table 11: The top 9 drug-gene interaction findings and their replication results from the UKBB primary care data.

** The results from the UKBB cohort is adjusted by age and sex.*

*** **(2)** is replicated, **(4)** is validated, and **(8)** and **(9)** being directionally consistent with borderline significance and **(1)** has a supporting evidence with another SNP (CYP2D6*4: poor metabolism) known to has an opposing effect to the one initially discovered (CYP2D6*2: normal metabolism).*

This page was left intentionally

3.2 The replication results for the top 3 drug-drug-gene findings from the UKBB primary care data.

Figure 17 at the end of DDGIs results section summarizes all results under this category.

In this section, none of the 3 associations has been replicated. However, I was able to validate our first association. Here, I present the results for all associations starting from the validated finding.

1) Atorvastatin-Bisoprolol-rs9516519 (T>G, *ABCC4*): increased risk of drugs' stopping.

The significance of this association was mainly driven by findings from GS and GoSHARE cohorts. GS participants on atorvastatin-bisoprolol combination who were G-allele carriers at rs9516519 were 15.84 (4.5-55.7) times more likely stop any of the two drugs during their combined used per allele ($p=1 \times 10^{-6}$). Similarly, the concomitant use of this combination among GoSHARE patients carrying this allele was associated with a 3.26 (1.12-9.51) times higher likelihood to stop either of the two drugs per allele (p=0.045). However, in UK Biobank, there was no association between the use of this drug combination and rs9516519 *ABCC4* variant (p=0.56).

As mentioned in the previous chapter, bisoprolol is primarily metabolized by CYP3A4, and atorvastatin is a CYP3A4 inhibitor. Accordingly, we examined the hypothesis that bisoprolol users co-treated with atorvastatin had a larger SBP drop than when bisoprolol was used alone (studying it as a drug-drug interaction (DDI)), and whether this response could increase with the presence rs9516519 *ABCC4* variant (studying it as a drug-drug-gene interaction (DDGI).

Two approaches were used: 1) comparing SBP levels before and after adding atorvastatin for bisoprolol users and 2) comparing SBP changes between two groups of patients (i.e. those on bisoprolol only vs those on atorvastatin + bisoprolol). For the first approach, 734 users of bisoprolol-atorvastatin combination from the UKBB cohort were identified. However, from these patients, we have only identified 69 patients who started bisoprolol first and then added atorvastatin; the rest (the majority, n= 665) of users started atorvastatin first and then added bisoprolol. The first group (n= 69) represents patients who are eligible for the first model as they have had a period while they were on bisoprolol alone before starting atorvastatin. Interestingly, examining the difference in SBP level for these 69 patients before and after adding atorvastatin using a linear mixed effect model provided some evidence supporting our proposed hypothesis: the mean SBP for bisoprolol users is reduced by ~ 3 mmHg after adding atorvastatin, p=0.062.

In the second method, a multiple regression model was utilized to study the difference in mean SBP levels between two groups of patients :1) all patients on bisoprolol alone (i.e. have no atorvastatin, n = 17,333) and 2) all patients who started atorvastatin first before adding bisoprolol (n= 665). Very interestingly, there was strong evidence supporting our hypothesis and confirming our first observation. I have observed a large difference in SBP levels between users of atorvastatin-bisoprolol combination and bisoprolol-only users. Atorvastatin users co-treated with bisoprolol experienced ~ 8 mmHg increased reduction in SBP as compared to the bisoprolol-only group ($p < 2 \times 10^{-16}$).

When the genotype variable for the rs9516519 (T>G) *ABCC4* SNP is added to the treatment group variable (i.e. on bisoprolol vs on the combination), a statically significant interaction has been observed between the two variables (p (interaction) = 0.02242).

When patients are stratified by their rs9516519 (T>G) *ABCC4* genotype, this variant was found to be linked with increasing the risk of SBP drop (p (trend) = 0.0059446). Carriers of the wild-type genotype (TT) and on this drug combination have shown ~ 7.4 mmHg lower mean SBP as compared to carriers of the same genotype but are on bisoprolol only. Those carrying the heterozygous genotype (TG), who are on the atorvastatin-bisoprolol combination, have experienced a mean SBP reduction by ~ 10 mmHg as compared to carriers of the same genotype from the bisoprolol-only group. This risk has increased even more with carriers of the recessive genotype (GG) as the mean SBP drop has reached to ~ 12 mmHg among the combination users compared to the bisoprolol-only group.

When results are stratified by treatment group, compared to the normal genotype (TT), the heterozygous (TG) and recessive (GG) genotypes have shown a clear trend toward increased SBP drop among the combination users (p (trend) = 0.02242) with the difference being - 2.46 and - 4.54 mmHg for both genotypes respectively.

2) Simvastatin-Metformin-rs622342 (A>C, *SLC22A1*): increased risk of drug stopping.

I have observed a consistent direction of effects across all cohorts with the results being significant for the GoDARTs cohort. Carrying the minor allele (C) at this variant in *SLC22A1* gene was associated with a 3.07 (1.31-7.2) , 3.66 (1.05-12.81) , and 2.4 (0.81-7.08) times increased chance of stopping simvastatin or metformin used in combination for GoDARTs (p=0.0074) , GS (p=0.057) , and GoSHARE (p=0.12) respectively . Nevertheless, I was not able to see a similar association within users of this combination from the UKBB primary care data (OR = 0.94, p=0.51).

3) Metformin-Gliclazide-rs1967120 (G>A, *ABCC1*): increased risk of dose decrease.

The minor allele (G) at this *ABCC1* variant was associated with a 19 % (p=0.0018), 19 % (p=0.54), and 26% (p=0.017) higher odds of decreasing the daily dose any of the two drugs when used in combination among GoDARTs , GS, and GoSHARE participants. However, we did not replicate this result in UK Biobank (OR = 1.04, p=0.72).

Figure 17 below summarizes the results of the top 3 DDGI associations from the combined Scottish cohort and their replication results from the UKBB primary care data, showing the steps used for the validation of the atorvastatin-bisoprolol-rs9516519 (ABCC4) association.

(1) Atorvastatin-Bisoprolol-rs9516519 (ABCC4). Drug stop phenotype		
(A) The discovery (the combined) cohort. *(controls = 2,571, cases = 18)*		
Cohort	**OR (CI)**	**P-value**
GoDARTs	0 (no cases)	1
GS	15.84 (4.5-55.7)	0.000001
GoSHARE	3.26 (1.12-9.51)	0.0456
Combined	4.72 (2.44-9.13)	0.0000148
(B) The replication (the UKBB) cohort. *(controls = 5,871, cases = 93)*		
UKBB*	0.96 (0.78-1.17)	0.56

(2) Simvastatin-Metformin-rs622342 (SLC22A1). Drug stop phenotype		
(A) The discovery (the combined) cohort. *(controls = 6,940, cases = 25)*		
Cohort	**OR (CI)**	**P-value**
GoDARTs	3.07 (1.31-7.2)	0.00747
GS	3.66 (1.05-12.81)	0.0577
GoSHARE	2.4 (0.81-7.08)	0.122
Combined	3.1 (1.71-5.62)	0.000093
(B) The replication (the UKBB) cohort. *(controls = 6,397, cases = 231)*		
UKBB*	0.94 (0.77-1.14)	0.51

(3) Metformin-Gliclazide-rs1967120 (ABCC1). Dose decrease phenotype		
(A) The discovery (the combined) cohort. *(controls = 4,333, cases = 1,302)*		
Cohort	**OR (CI)**	**P-value**
GoDARTs	1.19 (1.07-1.33)	0.001877
GS	1.19 (0.67-2.11)	0.5476
GoSHARE	1.26 (1.04-1.53)	0.0175
Combined	1.21 (1.1-1.33)	0.000093
(B) The replication (the UKBB) cohort. *(controls = 4,115, cases = 244)*		
UKBB*	1.04 (0.85-1.26)	0.7269

Is there a known route of interaction between the 2 drugs?

Yes

No

1

Further investigation using an alternative phenotype

Is adding atorvastatin (CYP3A4 inhibitor) to bisoprolol (CYP3A4 substrate) increases bisoprolol toxicity as indicated by increased systolic blood pressure (SBP) reduction?

Linear mixed effect model :-

Formula:
mean SBP after bisoprolol~ treatment context + mean SBP 1 year before bisoprolol + age + sex + (1 | eid)

Outcome (Y) variable:
Mean SBP **after** bisoprolol

(2 measurements per patient within 1 year : pre- and post- atorvastatin)

Mean SBP 1 year **before** bisoprolol | Treatment context : * Bisoprolol only OR * Bisoprolol + Atorvastatin | **Age** | **Sex**

(each patient has 2 mean SBP measurements taken within 1 year (one while on bisoprolol only and one after adding atorvastatin to bisoprolol))

Explanatory (X) variables

SBP level for bisoprolol users is reduced after adding atorvastatin.

Fixed effects	Estimate	Std. Error	t value	P-value
Intercept	98.91193	20.64621	4.791	9.96e-06 ***
Mean SBP 1 year before bisoprolol	0.28015	0.09863	2.840	0.00601 **
Bisoprolol + Atorvastatin *(vs. Bisoprolol only)*	**- 3.17453**	1.67852	-1.891	**0.06285** .
Age	-0.05320	0.24580	-0.216	0.82933
Sex (M) *vs. F*	1.93499	2.84619	0.680	0.49901

2

Further investigation for the above finding using a different model with a larger sample size

Multiple regression analysis :-

Formula:
*mean SBP **after** bisoprolol ~ mean SBP 1 year **before** Bisoprolol + patient's group ? + Age + Sex*

Outcome (Y) variable :
Mean SBP **after** Bisoprolol

(2 groups of patients :
a- On bisoprolol alone : mean SBP during 1 year on the drug.
b- Started atorvastatin before adding bisoprolol : mean SBP **during** the interaction time (i.e. after adding bisoprolol to atorvastatin and within 1 year of interaction))

Mean SBP 1 year **before** bisoprolol (i.e. before the bisoprolol for the 1st group and before the interaction time for the 2nd group) | Patient's group : **a)** Bisoprolol alone (n = 17,333) **b)** Atorvastatin + bisoprolol (n= 665) | Age | Sex

Explanatory (X) variables

Coefficients	Estimate	Std. Error	t value	P-value
Intercept	73.867570	1.114720	66.266	< 2e-16***
Mean SBP 1 year before treatment	0.401789	0.005281	76.087	< 2e-16***
Age	0.156945	0.014199	11.053	< 2e-16 ***
Sex (M) vs. F	0.452848	0.173940	2.603	0.0092 **
During Bisp. + Atorv. mean SBP *(vs.* Bisp. only SBP)	**- 8.106801**	0.459147	-17.656	**< 2e-16** ***

3

The influence of rs9516519 ABCC4 genotype on SBP level

(a)

Bisoprolol Only SBP						
Genotype	**n**	**mean**	**SE**	**dif**	**lower**	**upper**
A/A	12751	145.3	0.1178	0	NA	NA
A/a	4243	145.9	0.2095	0.2422	-0.1603	0.6448
a/a	339	145.1	0.6751	0.306	-1.0191	1.4803

Bisoprolol-Atorvastatin SBP						
Genotype	**n**	**mean**	**SE**	**dif**	**lower**	**upper**
A/A	500	138.3	0.6876	**-7.394**	-8.432	-6.356
A/a	153	137.4	1.2800	**-9.854**	-11.702	-8.006
a/a	12	134.8	4.2139	**-11.942**	-18.502	-5.383

p-value (interaction) = 0.02242

(b)

A/A						
Bisoprolol only SBP	12751	145.3	0.1178	0	NA	NA
interaction SBP	500	138.3	0.6876	**-7.394**	-8.432	-6.356
A/a						
Bisoprolol only SBP	4243	145.9	0.2095	0	NA	NA
interaction SBP	153	137.4	1.2800	**-10.1**	-11.97	-8.226
a/a						
Bisoprolol only SBP	339	145.1	0.6751	0	NA	NA
interaction SBP	12	134.8	4.2139	**-12.17**	-18.84	-5.502

p-value (trend) = 0.0059446

(c)

Bisoprolol only SBP						
	n	**me**	**se**	**dif**	**lower**	**upper**
A/A	12751	145.3	0.1178	0	NA	NA
A/a	4243	145.9	0.2095	0.2422	-0.1603	0.6448
a/a	339	145.1	0.6751	0.2306	-1.0192	1.4803
During interaction SBP						
A/A	500	138.3	0.6876	0	NA	NA
A/a	153	137.4	1.2800	**-2.460**	-4.558	-0.3609
a/a	12	134.8	4.2139	**-4.548**	-11.182	2.0857

p-value (trend) = 0.02242

Figure 17: The top 3 drug-drug-gene interaction findings and their replication results from the UKBB primary care data as adjusted by age and sex.

** The last 3 genetic models show the significance of interaction between genotype and treatment group **(a)**, the results as stratified by genotype **(b)**, and the results as stratified by treatment group **(c)**.*

This page was left intentionally blank

4. Discussion

4.1 The top 9 drug-gene interactions findings

Here, I discuss our findings, starting from the best evidence as ordered in the results section.

4.1.1 Replicated findings (n = 1).

Quinine-rs4918758 (CYP2C9)

Users of quinine from the combined cohort who carry the C allele at rs4918758 (T>C) SNP in the *CYP2C9* gene were found less likely to decrease their dose. Interestingly, this novel observation has been replicated among quinine users from the UKBB cohort as the same allele was found significantly linked with the same phenotype in a similar direction of effect. Quinine is extensively metabolized in the liver primarily by CYP3A4, but other enzymes including CYP2D6, CYP2E1, CYP2C8, CYP2C19, and CYP2C9 have been reported to be involved in quinine metabolism [246]. The C allele at rs4918758 *CYP2C9* SNP has been previously associated with decreased warfarin dosage requirements in Korean populations [247] suggesting that the variant might be associated with decreased enzyme activity. In addition, the CYP2C9 reduced activity variants *CYP2C9*2/*3*/11* were previously associated with decreased therapeutic efficacy of quinine derivatives chloroquine and primaquine in treating malaria [248]. Consistent with this, our results also show that both *CYP2C9*2* and *CYP2C9*3* variants are associated with decreased odds of reducing quinine dose. Taken together these results suggest that the observed quinine/CYP2C9 interaction with reduced efficacy or increased tolerability is associated with reduced function of CYP2C9. This seems unlikely to be an effect in quinine metabolism per se, which is largely metabolized by CYP3A4, but might reflect altered metabolism of downstream metabolites.

As discussed above, the interaction observed may not be mediated via a PK interaction. A recent study [249] reported that rs4918758 was associated with decreased coronary heart disease risk. Given that a main side effect of quinine is cardiac toxicity (i.e. prolongation of QT and arrhythmias), the drug-gene interaction could be explained by a reduction in cardiac toxicity with quinine in carriers of this cardioprotective variant. The importance of this novel finding lies in the fact this variant allele is very common among the British population with a minor allele frequency of ~37% as shown from the UKBB cohort. In addition, quinine is known by its undesirable and life-threatening side effects profile; and therefore, identifying patients' subgroups who might be well- tolerated to the drug could be a crucial clinical addition to prescribers.

4.1.2 Validated findings (n = 1).

Doxazosin-rs9895420 (ABCC3)

The combined cohort findings show that carriers of rs9895420 (T>A) mutation in ABCC3 transporter were at decreased risk to reduce their daily doxazosin treatment. Although I couldn't find this correlation to be significant from the UKBB primary care data, the best OR estimate was found consistent with the same direction of effect. Doxazosin is mainly metabolized in the liver, and 63% of the dose is excreted in the feces [250] suggesting a potential role of hepatic transporters in its elimination. It is unknown whether ABCC3 transporter contributes to the elimination process of doxazosin, but our findings show that the A allele at rs9895420 *ABCC3* variant could be associated with increased tolerability of the drug in term of side effects.

Interestingly, I was also able to uncover an association between this variant and increased doxazosin SBP lowering capacity. The mechanism by which this observation occurs is uncertain, but this variant has been previously linked with increased ABCC3 activity [251] suggestion its potential role to increase doxazosin efflux from the liver into systemic circulation leading to increased anti-hypertensive efficacy.

4.1.3 Findings with supporting evidence (n =1).

Ramipril-rs1135840 (CYP2D)

I have observed a strong correlation between rs1135840 (G>C) variant in *CYP2D6* gene and low likelihood of stopping ramipril treatment; an association which passed the nominal significance level (p <=0.05) and shows consistent directions of effects across all of the 3 cohorts formulating the combined cohort in which I find the association significant after the Bonferroni correction. However, and potentially due to genotyping issues related to this variant in the UKBB as it deviates from HWE, I couldn't find this association significant or with a similar direction of effect among the UKBB participants. Ramipril is a prodrug which undergoes renal and hepatic metabolism to be converted into its active metabolite ramiprilat [252]. 75% of the total ramipril metabolism occurs in the liver, with 25% catalyzed by esterases [252]. However, it is unknown whether CYP2D6 enzyme could contribute to the total ramipril metabolism. The minor allele (C) at the well-known rs1135840 (G>C) (*CYP2D6**2) variant represents the extensive (normal) metabolizer phenotype [253]. Assuming ramipril as a CYP2D6 substrate, the C allele could increase ramipril metabolism and increases its therapeutic efficacy consequently. This is consistent with our observation that this variant is linked with decreased chance to stop the treatment. Interestingly, our results also show that the other well-known loss-of-function variant, rs3892097 (C>T) (*CYP2D6**4) [253], is correlated with a 29% increased risk to stop ramipril treatment. Loss of the enzyme activity could inhibit

production of the active metabolites causing loss of efficacy which could lead to stopping of the drug. This other observation provides more evidence on a potential ramipril-CYP2D6 correlation. In fact, rs1135840 and rs3892097 are in LD to each other (D' = 1) with the C allele at rs1135840 correlated with the C allele at rs3892097. Further support of these findings can be also observed from our data as we find similar signals in the Scottish discovery cohort for both enalapril and lisinopril suggesting that our findings are consistent and applies to all ACE inhibitors. Interestingly, we also note a case report for a patient homozygous for the *CYP2D6*4* variant and discontinued ramipril therapy shortly after starting it due to ramipril-induced dry cough which is consistent with our finding that this variant is linked with increased likelihood of stopping ramipril [254].

4.1.4 Findings where their replications were not significant but directionally consistent with a p-value close to the significance level (n =2).

1) Amlodipine-rs868853(ABCC4)

The minor allele (C) at rs868853(C>T) *ABCC4* SNP was observed to be correlated with a large reduced chance to decrease the daily amlodipine dose within the combined cohort participants. The result from the UKBB cohort shows that the same SNP is linked with the same phenotype in a similar direction of effect, but this wasn't statistically significant. 90% of the drug is metabolized in the liver, with 10% of the parent compound and 60% of the metabolites to be renally excreted [255]. It is unknown whether amlodipine is an ABCC4 substrate, but a previous study has shown that the C allele is linked with reduced methotrexate plasma levels and increased ABCC4 activity [256]. Similarly, our findings show that this variant is associated with lower odds to decrease the dose suggesting decreased likelihood for side effects which

could be attributable to decreased amlodipine plasma levels. This suggested mechanism of interaction could occur if the site of interaction is in the kidney assuming amlodipine being an ABCC4 substrate.

2) Clopidogrel-rs12353214 (PTGS1)

I have also noted carriers of the rs12353214 (C>T) mutation in *PTGS1* gene to be less likely to stop clopidogrel therapy; an association which was directionally consistent although not significant in UKBB. PTGS1 (COX-1) is a known contributor in the regulation of the coagulation process by the production of prostaglandins which facilitate clotting formation [257]. Aspirin exerts its anticoagulant effect by inhibiting PTGS1 activity and is often co-prescribed with clopidogrel for enhanced anticoagulation efficacy. Similarly, genetic variants in *PTGS1* gene could influence the therapeutic efficacy of clopidogrel. Among clopidogrel users, the rs10306114 (A>G) and rs1330344 (T>C) variants in *PTGS1* gene have been previously reported to be linked with increased and decreased cardiovascular events (reduced and increased clopidogrel efficacy) for both SNPs respectively [258, 259]. In the same context, I have uncovered another novel and independent *PTGS1* variant, rs12353214 (C>T), as linked with lower odds to stop the treatment suggesting enhanced clopidogrel therapeutic efficacy.

4.1.5 Other non-replicated/non-validated associations (n = 4).

1) Nifedipine-rs152023 (ABCC1)

Although the minor allele (C) at rs152023 *ABCC1* variant was detected to be associated with increased likelihood to stop nifedipine treatment in the combined cohort, I couldn't see any evidence supporting this observation from the UKBB primary care data. Nifedipine undergoes hepatic metabolism by CYP3A4 enzyme and 60-80% of the drug is excreted in urine as inactive metabolites with the reminder eliminated via biliary excretion [260]. The ABCC1 transporter is expressed in the

basolateral and apical membrane of the liver and kidney, respectively. However, currently, there is no evidence that nifedipine is an ABCC1 substrate. In addition, the rs152023 SNP has not been reported previously to affect other drugs' pharmacokinetics [261].

2) Valproic acid-rs7916649 (CYP2C19)

Valproic acid users from the combined cohort who carry the A allele at rs7916649 (T>A) SNP in CYP2C19 enzyme have shown an increased tendency to reduce their daily dose. Nevertheless, this association was neither significant nor in the same direction of effect among the UKBB participants. The majority of the drug is eliminated via the liver where it is metabolized by multiple CYP enzymes, including CYP2C19 [262]. While there have been no previous studies linking the rs7916649 (T>A) variant with any drug response phenotypes, there is some evidence supporting the contribution of *CYP2C19* genotype in affecting valproic acid pharmacokinetics. The two well-known loss-of-function variants, *CYP2C19**2 and *CYP2C19**3, have been associated with increased valproic acid-induced weight gain in Japanese females and increased steady-state serum concentrations in Chinese patients [263,264].

3) Metformin-rs1045642 (ABCB1)

Carriers of the G allele at rs1045642 (A>G) *ABCB1* SNP from the combined cohort were less likely to stop metformin treatment which was also seen among the UKBB participants but wasn't statistically significant.

Metformin is mainly eliminated via renal excretion [265], and it has been suggested that metformin-induced gastrointestinal side effects could be linked to its intestinal accumulation due to the lack of its transportation, due to OCT1 reduced activity variants for example, from the intestine into systemic circulation [266]. Metformin has been shown to be transported by ABCB1 transporter [267]. The rs1045642 (A>G)

variant has been extensity studied and was correlated with multiple drug response phenotypes with mixed evidence regarding its linkage with decreased/increased ABCB1 expression [268]. The decreased ABCB1 intestinal expression could reduce intestinal metformin accumulation leading to reduced side effects and increased tolerability to the drug in term of side effects. On the other hand, reduced renal and intestinal ABCB1 activity could also result in increased metformin bioavailability leading to increased tolerability of the drug in term of its efficacy. These proposed mechanisms are consistent with our findings that carriers of this variant were less likely to stop the drug.

4) Nicorandil-rs5788 (PTGS1)

I have also detected a *PTGS1* variant, rs5788 (C>A), to be associated with an increased risk to stop nicorandil treatment among the combined cohort users. The best OR estimate from the UKBB results shows a similar trend of association with a modest effect size which wasn't statistically significant. Due to its vasodilatory effect, the angina treatment nicorandil is associated with hypotension-related side effects [269]. Prostaglandin synthas1 (PTGS1/COX1) has an important role in the conversion of arachidonic acid into prostaglandins which work as blood pressure regulators [270]. Therefore, genetic variants in this gene could influence nicorandil-induced hypotension. Our findings suggest that the A allele at rs5788 *PTGS* variant could augment nicorandil-induced hypotension as indicated by an increased tendency to stop the treatment.

4.2 The top 3 drug-drug-gene interactions findings

Here we discuss our 3 findings starting from the one which has been validated before discussing the other two associations which have not been replicated/validated.

1) Atorvastatin-Bisoprolol-rs9516519 (ABCC4)

To date, there is no pharmacokinetic interaction study on this combination. However, as clarified in the results section, atorvastatin is predicted to interact with bisoprolol via inhibiting its metabolizing enzyme: CYP3A4. This could result in increased bisoprolol efficacy/toxicity. Our finding from the combined cohort shows a large chance to stop any of the two drugs during their interaction for carriers of the G allele at rs9516519 (T>G) variant in ABCC4 transporter. Nevertheless, I couldn't replicate this result in the UKBB cohort using the same phenotype.

Attractively, I was able to uncover a novel drug-drug interaction between bisoprolol and atorvastatin. In a small sample of 69 patients, I found that there was a reduction in SBP when atorvastatin was prescribed to patients already on bisoprolol. Then, using a large sample of patients (~ 18K individuals), I was able to confirm this initial finding as the mean SBP was greatly lower among atorvastatin-bisoprolol users compared to bisoprolol-only users. Furthermore, I have then shown that carrying the rs9516519 (T>G) *ABCC4* variant is associated with increasing the likelihood of SBP drop among the users of this drug combination. The greater SBP drop in this treatment group was seen among carriers of the recessive genotype (GG) followed by the heterozygous genotype (TG) and then the wild-type genotype (TT). This variant has been previously correlated with decreased toxicity and plasma levels of methotrexate [271].

The authors suggested that this could have resulted from increased renal ABCC4 efflux activity since the kidney is the primary route for methotrexate elimination [272]. On the other hand, bisoprolol is eliminated equally via both the kidney and the liver [273]. The drug is also mainly metabolized in the liver by CYP3A4 suggesting the importance of hepatic transporters in bisoprolol pharmacokinetic. ABCC4 transporter is expressed in the hepatic basolateral membrane facilitating the efflux of xenobiotics into the systemic circulation. Our results from the combined cohort show that carriers of the rs9516519 (T>G) *ABCC4* variant are more likely to stop atorvastatin or bisoprolol when used concomitantly. This could result from increased bisoprolol toxicity which is what I was able to confirm in my DDGI analysis. The proposed mechanism for this drug-drug-gene interaction is that inhibiting CYP3A4 activity by atorvastatin in addition to increasing hepatic ABCC4 function by carrying the rs9516519 (T>G) variant could both result in a large bisoprolol systemic exposure and toxicity due to both decreased metabolism and increased efflux into the blood. This SNP is only significant with the use of both bisoprolol and atorvastatin but not with bisoprolol only according to our results from the DGIs study.

2) Simvastatin-Metformin-rs622342 (OCT1)

Our results from the combined cohort indicate that patients on this drug combination and carry rs622342 (A>C) variant in OCT1 transporter were more likely to stop any of the two drugs during their interaction. However, I couldn't replicate this association from the UKBB primary care data. Although OCT1 is a known transporter for metformin [274], currently, there is no known route of interaction between simvastatin and metformin. Nonetheless, in a previous study [275], the variant allele (C) has been observed to be associated with poor response to metformin which was

proposed to be resulting from reduced metformin uptake into the liver. On the other hand, our results from our drug-gene interactions study presented in chapter III show no significant correlation between this variant allele and stopping metformin treatment (p=0.39). Having said that, the association was only significant among metformin users co-treated with simvastatin but not metformin-only users. According to GeneATLAS, this variant was found to be highly correlated with hyperlipidaemia as well. Thus, it may be that this variant is interacting with statins and metformin via its effect on lipids rather than as a drug transporter.

3) Metformin-Gliclazide-rs1967120 (ABCC1)

I have found that carriage of the minor allele (G) at rs1967120 (G>A) SNP in ABCC1 transporter is correlated with higher chances to decrease the dose of any of the two drugs when used concomitantly. However, I was not able to see this association significant within users of this combination from the UKBB cohort. There is no known route of interaction between the two drugs, and this is a common safe combination usually prescribed for patients not well controlled with metformin alone. One recent report shows that metformin contributes to the reduction of ABCC1 transporter expression [276]. However, it is unclear why patients carrying the G allele at rs1967120 (G>A) SNP needed lower doses from either of the two agents. In fact, my results from the DGI study show a strong correlation between carriers of this variant allele and increased risk to decrease the daily dose of metformin (OR = 1.12, p=0.00075). The same trend but with a weaker correlation has also been observed among gliclazide users (OR = 1.09, p=0.06). These findings suggest that there might be a potential novel DG and/or DDGI within carriers of this variant allele occurring in an unknown mechanism of interaction which requires further investigation.

5. Conclusion

Out of 9 DGIs, 1 association was replicated (quinine-rs9418758 (CYP2C9)), one finding was validated (doxazosin-rs9895420 (ABCC3)), one finding has supporting evidence (Ramipril-*CYP2D6*2*), and two signals were directionally consistent with p-values close to the significance level (amlodipine-rs868853 (ABCC4) and clopidogrel-rs12353214 (PTGS1)). In addition, out of 3 DDGIs, one association was validated using an alternative phenotype but not replicated with the same phenotype (atorvastatin-bisoprolol-rs9516519 (ABCC4)). It is important to highlight that all results that didn't replicate or validate in this chapter should be interpreted with caution. I have classified all findings in this chapter into categories starting from higher to lower level of evidence depending on whether or not these have been replicated, validated, or have supporting evidence. This should provide some guidance on which results are more likely to be genuine. Regarding non-replicated/validated findings, these were included in the category of "lower level of evidence". We discussed these results only to highlight any potential clinical relevance to them which could worth further investigation but not to claim any definitive conclusions from them. One lesson I have learnt from our DDGI study is that it is really worth considering an alternative phenotype for replication if the same ‘generic’ phenotype did not replicate. In general, examining alternative phenotypes should always be encouraged even for the replicated findings as this will allow us to understand the novel discovered finding more deeply. For example, stopping the drug as a result of a certain DG/DDGI is a crucial clinical association since it could be linked with a change in the clinical practice. However, the exact reason for observing this phenotype will remain unknown until further phenotypes are examined.

Chapter VI:

Final General Discussion

1. Introduction

The field of pharmacogenomics has been growing rapidly in the last few years. This is reflected by the increasing number of publications each year. However, most of the work usually focuses on studying specific genotype-phenotype associations based on previous knowledge on the functional characteristics of the studied variant and the common elimination pathways of the studied drug. Nevertheless, our work has attempted not only to uncover novel pharmacogenomic associations but also potentially new and clinically significant genetic variants and/or novel drug targets. To do so, I have studied the associations between genetic variants in all important enzymes and transporters and their influence on commonly used drugs in the UK. I have covered the topic by considering both drug-gene and drug-drug-gene interactions. The latter is relatively a new topic in the field of pharmacogenomics with very limited publications to date. Therefore, in the first chapter, I established a detailed classification framework for these kinds of interactions by covering different mechanisms by which these interactions can occur based on observations from previous studies.

This thesis provides, for the first time, extensive coverage of clinical pharmacogenomic associations for 50 common chronic drugs and 50 commonly used chronic drug combinations in the UK. Studying the association of 162 genetic variants with 3 drug response phenotypes (drug-stop, dose-decrease, and genotype distribution changes among drug users) for each drug and each drug combination has generated a total of 48,600 results divided equally between the two studies (DG and DDGI) and are accessible via 2 online databases. I have utilized two cohorts with different nature: the UKBB cross-sectional prescribing data and the combined cohort longitudinal prescribing data for our discovery findings (i.e., the 48,600 findings).

Then, I have made further investigations for our top findings from the combined cohort by utilizing another larger and independent longitudinal cohort based upon the UKBB primary care data.

The main findings from this PhD project can be classified into three categories: novel associations from the combined cohort which have been replicated, validated, or were directionally consistent with p-values close to but not reaching significance level in the UKBB primary care data (n=5); associations consistent with previously reported findings (n = 19); and associations which are significant after Bonferroni correction from the UKBB cross-sectional cohort but require further investigation when larger cohorts become available (n = 11). In the sections that follow, I discuss the first category in more details with a brief review of the other two categories, highlighting points of strength and weakness of the research, and finally, I discuss my future directions.

2. The main PhD project findings

2.1 Novel associations which have been replicated, validated, or were directionally consistent with p-values close to significance (n=5).

2.1.1 Quinine-rs4918758 (T>C) *CYP2C9* variant [replicated]

Quinine is an FDA-approved drug for the treatment of uncomplicated *Plasmodium falciparum* malaria [277]. However, its use is challenged by concerns about its safety profile and the presence of safer and more efficacious agents. The recent evidence suggests that intravenous artesunate is the drug of choice in the treatment of severe malaria with quinine being an alternative agent [278]. In addition, artemisinin-based combination therapy (ACT), which consists of artemisinin derivatives combined with another anti-malaria agent from a different class such as lumefantrine, mefloquine or

amodiaquine, now represents the first-line option for the treatment of uncomplicated malaria [278]. However, quinine is still widely used for the management of uncomplicated malaria due to the limited availability of ACT.

In malaria-free countries, quinine is commonly used for the treatment of nocturnal leg cramps (NLC) which are painful chronic muscle strains usually occur at bedtime causing severe disruption of sleep. As the exact cause of these cramps is unknown, to date, there is no proven treatment for this condition [279]. In fact, quinine is used as off-label treatment for this condition and it is not approved by the FDA as risks of taking the treatment outweigh its modest therapeutic efficacy against nocturnal leg cramps [280]. The FDA's Adverse Event Reporting System (AERS) (2005-2008), shows 38 cases of serious adverse events from quinine such as hearing loss, rash, electrolyte disturbances, thrombocytopenia (the most common side effect), mucosal bleeding, and even two cases of death [280]. Despite this FDA warning, quinine is still commonly prescribed in the UK for nocturnal leg cramps as indicated by our findings from the UKBB prescribing data that shows that quinine is ranked 30th within the list of most commonly used chronic drugs in the UK. However, its usage in the UK is not recommended as a routine treatment for NLC but is only allowed if the cramps are very painful and continuous, other causes of cramps which can be treated have been excluded, and other non-pharmacological options are not working [281].

These adverse events restricting the use of the drug could be influenced by the individual variability in genes controlling the pharmacokinetics of quinine. In two case reports [282,283], individual differences can affect quinine plasma levels and therapeutic efficacy. In spite of the adequate dosing in the two cases, quinine plasma levels were abnormally low, leading to the death of the first case as a result of malaria exacerbation. As outlined earlier in chapter III, quinine is mainly metabolized by

CYP3A4 but other enzymes, including CYP2C9, are also involved in quinine metabolism. In one clinical trial, ciprofloxacin (a CYP3A4 inhibitor) has been shown to significantly increase quinine plasma levels suggesting increased quinine side effects, and therefore dose reduction could be recommended [284]. Regarding CYP2C9, no previous studies were found connecting CYP2C9 and quinine safety/efficacy. Nevertheless, in chapter V, I have presented a study showing that reduced CYP2C9 activity variants were linked with decreased chloroquine/primaquine treatments efficacy against malaria. Here, I have discovered for the first time a novel and replicated association between the rs4918758 (T>C) genetic variant in the *CYP2C9* gene and quinine daily dose. Individuals carrying this variant allele had a lower tendency to decrease their quinine dose. This could reflect that this subgroup of patients may have lower side effects and/or reduced efficacy. The latter phenotype could provide some potential explanation for the treatment failure seen in the two case reports mentioned above. Interestingly, the variant allele (C) at this SNP is very common in the UK with a frequency of ~ 37% in the UKBB cohort. This variant is also generally very common across all ethnicities with distribution frequencies of ~ 30%, 25.5%, ~39.4%, 36.4%, and 49.8% among Africans, Americans, East Asians, Europeans, and South Asians respectively [285]. Therefore, if future clinical or in-vitro studies confirm our finding, there may be a large clinical impact on quinine prescribing. Based on the rs4918758 (T>C) *CYP2C9* variant, patients might be stratified into two groups: those at a decreased risk for quinine toxicity but also could have a higher chance for low efficacy and those who are with expected increased toxicity and efficacy. This stratification would be helpful to guide the prescriber which groups of patients are at higher risk for toxicity, and therefore a lower dose could be prescribed. This is helpful when it comes to

prescribing the drug for leg cramps since toxicity was the main concern raised by the FDA. Regarding prescribing the drug for malaria, identifying patients who might experience lower efficacy is very helpful to protect patients from death caused by malaria exacerbation due to poor quinine response.

2.1.2 Doxazosin-rs9895420 (T>A) *ABCC3* variant [validated]

Doxazosin is an alpha-1 receptor antagonist indicated for the treatment of benign prostatic hyperplasia, but it is more commonly used for the treatment of hypertension. Inhibiting alpha-1 receptor results in vasodilation of arteries and veins, which in turn, results in decreasing peripheral resistance and blood pressure [286]. The current treatment guidelines for hypertension recommend using alpha-blockers as an add-on treatment for patients with resistant hypertension who are not well controlled, even with three different classes of antihypertensive agents [287].

The findings from the combined Scottish cohorts indicate that carriage of the A allele at rs9895420 (T>A) SNP in the ABCC3 transporter is associated with decreased odds of dose reduction of doxazosin which was directionally supported by the UKBB primary care analysis. This evidence of a potential rs9895420 (ABCC3)-doxazosin correlation has led us to study the effect of this SNP on the SBP reduction achieved by doxazosin. Excitingly we were showed that this SNP was significantly linked with increased antihypertensive efficacy of doxazosin. The exact mechanism by which this drug-gene interaction has occurred needs further investigation, but as outlined in chapter III, this could be resulting from increased doxazosin efflux into systemic circulation as this variant was reported to be associated with increased hepatic ABCC3 transporter activity if doxazosin was uncovered to be an ABCC3 substrate.

As doxazosin is mainly prescribed for difficult to manage patients, data on which patient subgroups could achieve better hypertension management could be of great clinical importance. The frequency of the variant allele (A) is 10% among the UKBB cohort participants (according to our findings) and ~ 12 % in Europeans [288]. It represents 6.20 %,11.25 %, and 8.43% of the American, South Asians, and East Asians populations respectively with its highest distribution being in Africans (~21%) [288] suggesting that our finding could be most relevant for patients from this ethnicity.

2.1.3 Amlodipine-rs868853 (C(*minor allele*)>T) *ABCC4* variant [directionally consistent with a p-value close to the significance level]

Hypertension management includes the use of four common drug classes: thiazides diuretics, angiotensin-converting enzymes inhibitors (ACEIs), angiotensin receptor blockers (ARBs), and calcium channel blockers (CCBs). The latter class is classified into two categories: non-dihydropyridines (non-DHPs) CCBs, such as verapamil and diltiazem, and dihydropyridines (DHPs) CCBs such as nifedipine, felodipine and, amlodipine. In addition to its usage in the management of chronic stable angina, amlodipine has been recently (2016) suggested to be considered the first-line agent in the management of hypertension compared to antihypertensive agents of other classes [289,290]. This drug is very commonly used in the UK as it was the 5th commonly prescribed chronic medication among the UKBB cohort participants. In addition to its once-daily tablet direction which increases adherence to the drug, amlodipine is also highly effective in the management of angina and hypertension with compelling evidence that the drug is associated with a strong reduction in cardiovascular events [290].

However, the use of CCBs could be hindered by their common and troubling side effect, which is peripheral oedema (PO). The incidence of peripheral oedema has been observed to be significantly higher among CCB users compared to the placebo group (10.7% vs 3.2%) with the discontinuation percentage being also significantly higher among CCBs users (2.1% vs 0.5%) [291]. The incidence of peripheral oedema was also seen significantly higher among DHPs CCBs (including amlodipine) users compared to non-DHPs CCBs users (12.3% vs 3.1%) [291]. Amlodipine side effects and efficacy could be influenced by genetic variability in metabolizing enzymes or transporters. As clarified previously, amlodipine is mainly metabolized by CYP3A4. Genetic variants in this gene have been shown to affect amlodipine blood-pressure-lowering efficacy in African-American population [292]. It has also been recently (2018) reported that carriers of the minor allele (C) at rs1045642 (C>T) SNP in the ABCB1 transporter experienced low amlodipine efficacy and increased side effects risk [293]. Our study has also uncovered another novel drug transporter may affect amlodipine tolerability. I have found that carriage of the minor allele (C) at rs868853 (C >T) mutation in the ABCC4 transporter is associated with being less likely to decrease amlodipine daily dose; an association which was nearly replicated in a larger independent cohort. The C allele frequency is ~ 7% of the UK population (according to our results from the UKBB cohort) with a similar percentage in Europeans [294]. Of note, the highest frequency of this minor allele is seen among Africans where it represents ~ 37% of the population with the frequencies being 6.7 %, 16.2%, and 10.8% among American, East Asian, and South Asian ethnicities respectively [294].

The lower tendency to reduce amlodipine dose among carriers of this variant allele could indicate reduced side effects. If these findings are confirmed in future studies, use of this genotype in prescribing decisions could reduce the incidence of amlodipine-induced side effects especially peripheral oedema; the most annoying adverse event of this very common drug in the treatment of hypertension. Patients can be stratified by their rs868853 (C >T) *ABCC4* genotype into increased or decreased risk groups to develop side effects so amlodipine dose could be decreased, or the drug can be switched into an alternative anti-hypertensive agent for the first group while the drug can be safely prescribed for the second group. Of note, the minor allele (C) is highly distributed in Africans, and therefore they may be the most affected ethnic group by our finding. Currently, there are no clinical trials comparing safety/efficacy profiles of amlodipine in different ethnic groups. Trials of this kind could help in understanding our results. Our findings suggest that the variant allele (C) could be linked with reduced side effects or efficacy of amlodipine. Therefore, if Africans (who have a greater frequency of the C-allele) were found to be more likely to develop any of these two phenotypes compared to other ethnic groups, this will provide more support to our finding. In addition, our result could also raise the attention into a potential novel amlodipine transporter (ABCC4) or could lead to the discovery of new mechanisms of interactions between amlodipine and ABCC4 transporter.

2.1.4 Clopidogrel-rs12353214 (C>T) *PTGS1* variant [directionally consistent with a p-value close to the significance level]

According to the World Health Organization (WHO), cardiovascular diseases (CVDs) represent the top leading cause of death globally, with ~ 18 million lives lost annually [295]. Decreased blood supply is mainly caused by the development of clots in blood vessels which hinder normal blood flow. Antiplatelet agents are one of

the most common therapeutic options to prevent thrombus formation by inhibiting platelet aggregation. Clopidogrel and aspirin are the most frequently prescribed antiplatelet drugs; each of which works on a different pathway to inhibit clot formation. Clopidogrel active metabolites inhibit P2RY12 receptor on the platelet and prevents binding of adenosine diphosphate (ADP) with the receptor. This leads to a series of inhibitions of other molecules ending by inhibiting fibrinogen formation which, in turn, results in inhibiting platelet activation. On the other hand, aspirin works by inhibiting the PTGS1 (COX-1) enzyme which initiates a series of inhibitions of other molecules that differ from those seen with clopidogrel inhibition but end by inhibiting the same molecules as with clopidogrel leading to a similar outcome. Figure 18 below (extracted from reference 296) summarizes mechanisms of actions of both drugs.

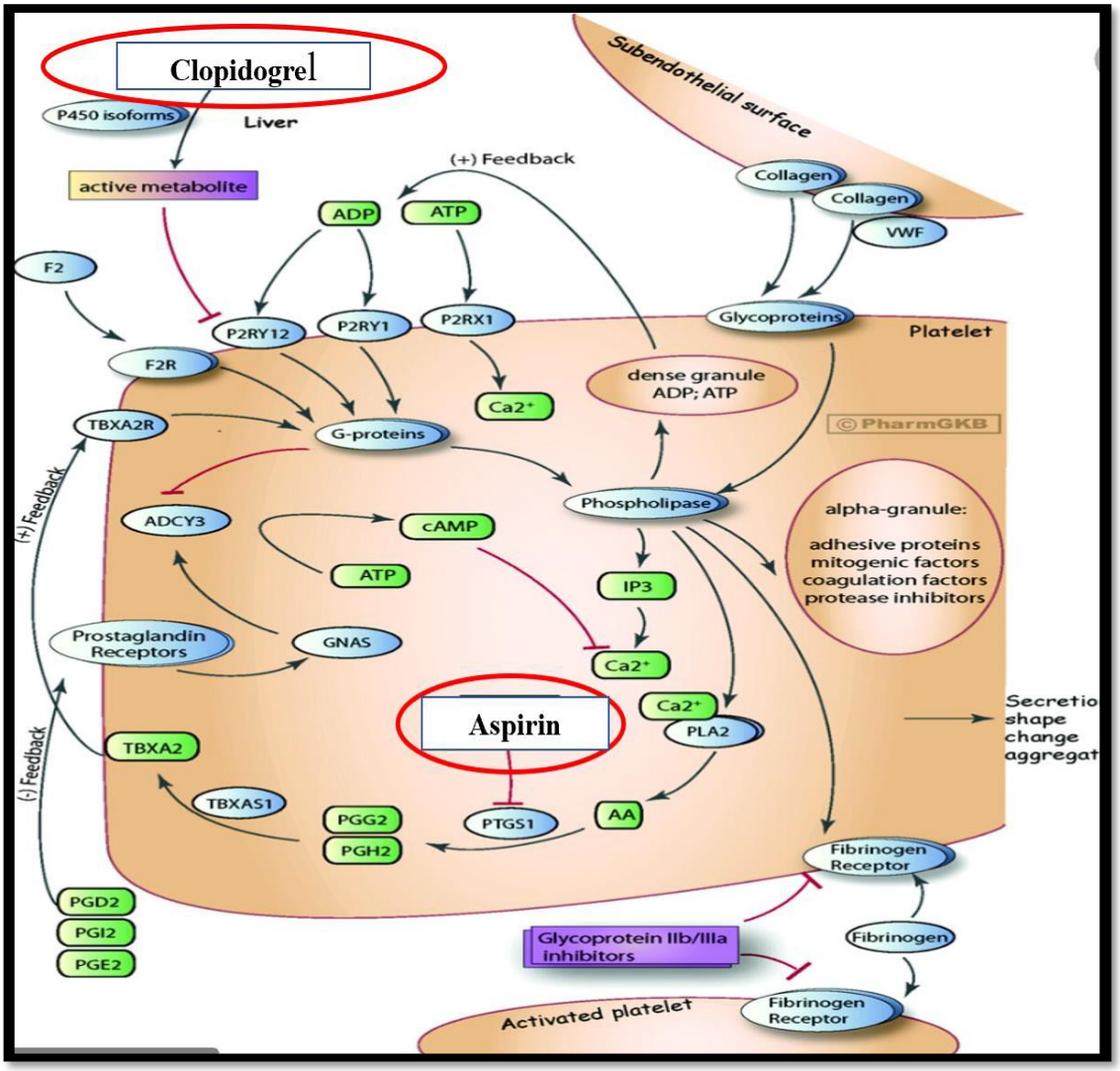


Figure 18: The different inhibitory pathways for both clopidogrel and aspirin which lead to the same pharmacological action (inhibiting platelet aggregation).

* *This figure has been reproduced from PharmGKB.*

In one study [297], clopidogrel was shown to be as effective as aspirin in the prevention of stroke and myocardial infarction. However, a previous clinical trial found that clopidogrel was superior to aspirin in reducing stroke and myocardial infarction [298]. Nevertheless, clopidogrel resistance is not an uncommon issue in clinical practice. The data from 10 previous studies show that the percentage of clopidogrel resistance among users of the drug is generally more than 20% with a minimum percentage being 5% and a maximum percentage reaching up to 44% [299]. Antiplatelet drug resistance could be defined by either development of atherothrombotic episodes during the treatment period or lack of antiplatelet-induced inhibitory effect on platelet aggregation as tested in the laboratory [300]. Resistance to antiplatelets could occur as a result of many factors such as lack of compliance, underdosing, and drug interactions. Antiplatelet resistance could also be related to platelet function as it could occur as a result of increased rate of platelet production or increased sensitivity of platelets to ADP. In fact, it has been previously reported that clopidogrel resistance is associated with the variability in platelet response to ADP between different individuals and that administration of clopidogrel did not increase this variability [301]. This suggests that clopidogrel resistance might not be only linked to the drug itself but with the presence of pre-existing factors. Aspirin could be also affected by these factors in addition to the variability in PTGS1 activities between individuals.

As aspirin and clopidogrel work in a complementary way on different pathways to inhibit platelet aggregation, the combination therapy of both agents has been considered the gold standard treatment option for ACS in the recent years [302,303]. Interestingly, I have found a novel genetic variant (rs12353214 (C>T)) in the *PTGS1* gene which seems to affect clopidogrel prescribing behaviour; carriers of the variant

allele (T) on this SNP were less likely to stop clopidogrel treatment. There was a similar association in UK Biobank which was close to, but not reaching, the significance level. Attractively, the recessive genotype (TT) at this variant has been previously reported to be associated with increased risk of myocardial infarction (MI) and unstable angina among non-steroidal anti-inflammatory drugs (NSAIDs) users [304]. As clopidogrel is indicated for the management of these two conditions, which seemed to be triggered furthermore by carrying this variant allele, it is more likely that carriers of this variant to show more adherence to this treatment due to a potential increased risk of CHD events. This is consistent with our finding that carrying of this variant was associated with decreased likelihood to stop clopidogrel treatment. However, currently, there are no studies to show whether or not this variant alone is directly linked with increased MI risk; the current available evidence only shows an association among NSAIDs users. More research is required to validate our finding.

The frequency of the variant allele (T) is 10% among the UKBB cohort participants and ~ 11% in Europeans [305]. It is very rare among Africans (0.91%), represents 9.22% and 15.8% of the American and South Asians populations respectively, and is highly distributed among East Asians (~ 41%) [305]. The fact that a considerable percentage of patients could fail to achieve therapeutic targets with clopidogrel makes it crucial to identify those at increased or decreased risk in order to prevent further deaths caused by ACSs. There are many factors affecting clopidogrel efficacy including genetic variants in the ABCB1 transporter and CYP2C19 and CES1 enzymes [306]. In addition, in chapter V, I presented two studies linking genetic variability in PTGS1 gene to clopidogrel efficacy. Our findings show another novel genetic variant (rs12353214 (C>T)) in PTGS1 gene associated with decreased risk of stopping clopidogrel therapy; an association which could mostly affect East Asians due to increased variant distribution.

Confirming this in future clinical studies will provide another piece of information to prescribers to enable more precise management of ACS or stroke with clopidogrel. Patients could be divided according to their *CYP2C19*, *CES1*, and/or *PTGS1* genetic profiles into responders or non-responders and/or those at increased/decreased risk for clopidogrel-induced bleeding enabling targeted, genotype determined dosing or use of alternative antiplatelet agents such as Ticagrelor.

2.1.5 Bisoprolol-Atorvastatin-rs9516519 (T>G) *ABCC4* variant [validated].

Bisoprolol belongs to the family of selective beta-1-blockers, which also include atenolol, metoprolol, and nebivolol. Blocking beta-1-receptor action results in decreased heart contractility, heart rate, cardiac output, and norepinephrine/renin levels [307]. Therefore, beta-1-blockers are indicated for the management of hypertension (by decreasing cardiac output and renin levels [308]), heart failure (by reducing the adverse effects of norepinephrine on myocytes [309]), post-myocardial infarction and chronic stable angina (by reducing oxygen demand as a result of decreased heart rate and contractility [310,311]). Selective beta-1-blockers are now considered the first-line agents for the management of chronic unstable angina [312].

The most common side effects of selective beta-1-blockers including bisoprolol are hypotension, bradycardia, low exercise capacity, nausea, vomiting, dizziness, headache, fatigue, and dry mouth/eyes [313]. Another risky side effect, which could mostly affect diabetic patients on glucose-lowering agents, is masking of the symptoms of hypoglycemia such as tachycardia which could delay the response to treat the condition leading to more complications [313].

The cases of toxicity from bisoprolol which have been documented to date presented mostly with bradycardia and/or hypotension [314]. The chance of this toxicity increases with the presence of other factors such as drug-drug and drug-gene interactions. As outlined before in chapter V, bisoprolol is mainly metabolized by CYP3A4, implying that CYP3A4 inhibitors could increase the risk of bisoprolol-induced toxicity. Our initial (discovery) finding shows that bisoprolol users co-treated with atorvastatin (a CYP3A4 inhibitor) who carry the rs9516519 (T>G) variant in the ABCC4 transporter were at high risk to stop either of the two drugs during their combined use. In the replication cohort, I was not able to see the same association with the same phenotype. However, when I examined the hypothesis that the two potential risk factors (atorvastatin + rs9516519 (T>G) *ABCC4* SNP) could increase bisoprolol-induced toxicity as indicated by greater systolic blood pressure (SBP) lowering, large drug-drug and drug-drug-gene interactions were observed. The most prominent effect seen was the drop in SBP by ~ 8 mmHg among bisoprolol-atorvastatin users compared to bisoprolol-only users. This drop in SBP was even greater amongst carriers of the rs9516519 (T>G) ABCC4 variant allele. This variant is mostly distributed among Americans (18.30%) and Europeans (13.22%) [315]; with the percentage being ~14% in the UK (according to our results from chapter III) while it is rarely found among South Asians (2.8%), East Asians (0.40%), and Africans (0.30%) [315]. It could be the case that the reason for stopping one of the two drugs during their interaction for carriers of this variant allele is explained by increased bisoprolol-induced toxicity. A drop in SBP by >= 8 mmHg could be clinically significant and lead to the development of symptoms of hypotension such as dizziness, nausea, fainting, fatigue, loss of concentration, blurry vision, and tachycardia. The chance of this risk could increase with non-hypertensive patients

who are prescribed bisoprolol for the management of other conditions, rather than hypertension, such as heart failure, myocardial infarction, or chronic stable angina.

Our findings provide the first evidence for a potential harmful drug-drug interaction between bisoprolol and atorvastatin and also shows potentially novel genetic variant (rs9516519 (T>G)) and transporter (ABCC4) which seemed to affect this drug-drug interaction. Further clinical trials and/or in-vitro studies are required to confirm these observations. If these findings are confirmed, then the prescribing guidelines for bisoprolol could change in the future. For patients already on atorvastatin or those who will add atorvastatin, bisoprolol dose could be reduced, or the drug could be switched into an alternative beta-1 blocker not influenced by the CYP3A4 enzyme such as atenolol. If the patient carries the G allele at rs9516519 *ABCC4* SNP, the recommendation will more likely be switching the drug as the risk of bisoprolol toxicity is expected to be higher for this subgroup of patients. As the distribution of this variant allele is higher among White Americans and White Europeans, the application of our DDGI finding could be of more interest among these two ethnic groups.

2.2 Associations consistent with previous studies' findings.

For drug-drug-gene interactions, all of our findings were novel, so we cannot identify whether any of our results replicated previously reported findings. However, for the drug-gene interactions study, I was able to identify 19 results from both the combined and the UKBB-cross-sectional cohorts to be consistent with previously reported findings. I have discussed these findings in chapter III. In Table 13 below, I summarize these findings.

No.	Drug	SNP_id	Gene	Our study conclusion	The other study/ies findings
1	Gliclazide	rs1057910 (A>C) (*2C9*3*)	*CYP2C9*	Increased likelihood for dose reduction > increased toxicity.	Increased risk for gliclazide-induced hypoglycemia.
2	Simvastatin	rs2231142 (G>T)	*ABCG2*	Decreased likelihood for stopping of the drug > increased efficacy or decreased side effects.	Increased simvastatin lactone plasma concentration and decreased clearance > increased chance of increased hepatic concentration > increased efficacy
3	Methotrexate	rs9895420 (T>A)	*ABCC3*	Decreased likelihood for stopping of the drug > increased efficacy or decreased side effects.	Reduced thrombocytopenia side effect/ reduced efficacy.
4	Methotrexate	rs1128503 (A(*minor allele*)>G)	*ABCB1*	The A allele is linked with an increased likelihood for stopping the drug > increased side effects or decreased efficacy.	The A allele is linked with increased risk of methotrexate toxicity
5	Carbamazepine	rs762551 (C(*minor*)>A)	*CYP1A2*	The C allele is linked with dose decrease > increased side effects	The C allele is linked with decreased clearance > increased plasma concentrations and side effects
6	Carbamazepine	rs2242480 (C>T)	*CYP3A4*	Decreased likelihood for dose reduction > increased likelihood of increasing the dose.	Decreased plasma concentration > low efficacy > dose increase.
7	Carbamazepine	rs1128503 (A(*minor allele*)>G)	*ABCB1*	The A allele is linked with increased likelihood for stopping the drug > increased side effects or decreased efficacy.	The A allele is linked with increased clearance > decreased efficacy
8	Pioglitazone	rs10509681 (T>C) (*2C8*3*)	*CYP2C8*	Increased likelihood for stopping the drug > increased side effects or decreased efficacy.	Decreased plasma concentration > decreased efficacy
9	Warfarin	rs2242480 (C>T)	*CYP3A4*	Increased likelihood for stopping the drug > increased side effects or decreased efficacy.	Increased clearance > decreased efficacy
10	Atorvastatin	rs2032582 (A(*minor allele*)>T)	*ABCB1*	The A allele is linked with decreased likelihood for stopping of the drug > increased efficacy or decreased side effects.	The A allele is linked with increased efficacy
11	Clopidogrel	rs1057910 (A>C) (*2C9*3*)	*CYP2C9*	Decreased likelihood for dose reduction > Decreased side effects/efficacy	Decreased efficacy
12	Amitriptyline	rs1065852 (G>A) (*2D6*10*)	*CYP2D6*	Low distribution of the variant allele carriers > increased side effects or low efficacy	Increased nortriptyline plasma level > increased side effects
13	Amlodipine	rs1045642 (A>G)	*ABCB1*	Low distribution of the variant allele carriers > increased side effects or low efficacy	Low clearance/high AUC > increased side effects
14	Simvastatin	rs4149056 (T>C) (*1B1*5*)	*SLCO1B1*	Low distribution of the variant allele carriers > increased side effects or low efficacy	Increased side effects (level 1A evidence)
15	Clopidogrel	rs12248560 (C>T) (*2C19*17*)	*CYP2C19*	High distribution of the variant allele carriers > low side effects or high efficacy	Decreased cardiovascular events/ increased drug activation > increased efficacy (level A evidence)
16	Clopidogrel	rs4244285 (*2C19*2*)	*CYP2C19*	Low distribution of the variant allele carriers > increased side effects or low efficacy	Low efficacy (Level 1A evidence)
17	Citalopram	rs28371725 (C>T) (*2D6*41*)	*CYP2D6*	High distribution of the variant allele carriers > low side effects or high efficacy	Increased efficacy
18	Pioglitazone	rs10509681 (T>C) (*2C8*3*)	*CYP2C8*	Low distribution of the variant allele carriers > increased side effects or low efficacy	Decreased AUC > Low efficacy
19	Fluoxetine	rs1065852 (G>A) (*2D6*10*)	*CYP2D6*	Low distribution of the variant allele carriers > increased side effects or low efficacy	Increased AUC > increased side effects

Table 13: Summary of 19 drug-gene associations consistent with previous studies' findings.

2.3 Associations require further investigation in a larger cohort.

I have found 11 novel DG/DDG associations from the UKBB cross-sectional prescribing data which were significant after Bonferroni correction. However, I was unable to check for the replicability of these findings as no ideal replication data was available to us. The ideal replication data should be another independent and larger cross-sectional prescribing data. In chapters III and IV, I have already discussed these findings. Here, I briefly outline these 11 findings in Table 14 below.

No.	Drug/drug combination	SNP_id	Gene	The associated phenotype
1	Lansoprazole	rs555754	*SLC22A3*	Low distribution of the variant allele carriers > increased side effects or low efficacy
2	Bendroflumethiazide	rs3743527	*ABCC1*	Low distribution of the variant allele carriers > increased side effects or low efficacy
3	Gabapentin	rs8187843	*ABCC1*	Low distribution of the variant allele carriers > increased side effects or low efficacy
4	Rosuvastatin	rs2231135	*ABCG2*	High distribution of the variant allele carriers > low side effects or high efficacy
5	Simvastatin-Furosemide	rs4148739	*ABCB1*	High distribution of the variant allele carriers > low side effects or high efficacy
6	Atorvastatin-Metformin	rs10937158	*ABCC5*	High distribution of the variant allele carriers > low side effects or high efficacy
7	Amlodipine-Atorvastatin	rs3735451	*CYP3A4*	High distribution of the variant allele carriers > low side effects or high efficacy
8	Bendroflumethiazide-Metformin	rs2199939	*ABCG2*	Low distribution of the variant allele carriers > increased side effects or low efficacy
9	Atorvastatin-Metformin	rs2293001	*ABCC5*	Low distribution of the variant allele carriers > increased side effects or low efficacy
10	Atorvastatin-Metformin	rs17731538	*ABCG2*	Low distribution of the variant allele carriers > increased side effects or low efficacy
11	Simvastatin-Metformin	rs215095	*ABCC1*	Low distribution of the variant allele carriers > increased side effects or low efficacy

Table 14: Summary of 11 drug-gene associations require further investigation in a larger cohort.

3. Points of strength and weakness of our research

I started this PhD project at the time when the UKBB prescribing data was available only as cross-sectional data in 2017. Therefore, the only way we found to define drug response phenotypes in this kind of data was by observing changes in the genotype distribution between drug/drug combination users and non-users. For our DGI study, the challenge of this approach lies in the fact that this phenotype could detect not only genetic variants affecting drug response but also variants which might be associated with diseases for which these drugs are prescribed for. To help minimize this risk, we excluded any drug-variant association where the variant was most likely associated with diseases but not drug response by utilizing the GeneAtlas database which provides millions of variants-traits associations using the UKBB cohort. In addition, DDGIs phenotypes cannot be accurately defined in a cross-sectional database due to uncertainty regarding the overlapping period between the two prescribed drugs and therefore the detected association might not necessarily reflect a DDGI, but could more simply reflect a DGI. To account for this issue, all DDGI findings in chapter IV were discussed considering DGI results from chapter III. So, for any DDGI, I also mention the results for each drug individually with the same variant, so we recognize whether our observation is a DDGI, DGI, or both.

Having mentioned these limitations in the UKBB cross-sectional data, one should not ignore its great power in detecting previously unrecognized associations giving the huge cohort size. With around 500,000 participants, the UKBB cross-sectional data represents the world's largest genetic database available to date.

In contrast to the UKBB cross-sectional data, I have also utilized three Scottish cohorts with longitudinal prescribing which provide a relatively small cohort size, even after combined together, compared to the UKBB. However, the advantage of this cohort is that it enabled us to define more accurate drug response phenotypes such as drug-stop and dose-decrease among users of the same drug/drug combination. For DDGIs, the longitudinal data enabled us to identify the exact interaction period between each of the two drugs. However, for the DDGI analysis, the numbers of patients prescribed two drugs concurrently who have a particular genotype is usually very small, which limits the power for these studies in the Scottish cohorts.

The points of strength of this project are that it covers for the first time a large variety of drug-gene and drug-drug-gene associations, utilized 4 different UK cohorts, and shows different methodologies for defining drug response phenotypes including both general and specific phenotypes in both cross-sectional and longitudinal prescribing data. Moreover, the majority of our findings from our DGI study were novel with almost all findings from our DDGI study to be novel giving the fact that this is the first work studying this topic comprehensively; however, clearly, further replication of all results is required before these results can be taken forward clinically. One other feature of this work is that it provides two user-friendly online applications to view all 48,600 results I generated in this project for both DG and DDGI studies.

We can gain considerable confidence in the findings we report because a total of 19 results from all different phenotypes ('drug-stop', 'dose-decrease', 'genotype distribution changes') were validated or replicated previous findings in some way. For example, validating findings from previous studies including the association between *CYP2C9*3/8*3* variants and increased risk of side effects/inefficacy from gliclazide/pioglitazone treatments. We have also validated the finding that loss-of-

function variants in ABC transporters family are linked with increased statin efficacy. What is more, our study has also validated some findings currently classified as 'level A evidence' (i.e., supported with strong evidence) from PharmGKB including the common associations between *SLCO1B1*5* variant and increased simvastatin side effects/intolerance, *CYP2C19*17* and increased clopidogrel efficacy/tolerance, and *CYP2C19*2* and decreased clopidogrel efficacy/tolerance. In addition, unlike many other publications in pharmacogenomic, our work also shows the replication results of the top discovery findings by utilizing a larger and independent UK cohort: the UKBB primary care data (recently available on September 2019 on ~230,000 participants).

Out of 8 DG associations (excluding one that cannot be replicated due to genotyping issues), 6 associations (75%) were found directionally consistent in the replication cohort. Four of these associations were of more interest. The Quinine-rs4918758 (CYP2C9) association was clearly replicated, doxazosin-rs9895420 (ABCC3) association was validated, and the amlodipine-rs868853 (ABCC4) and clopidogrel-rs12353214 (PTGS1) associations were directionally consistent with p-values close to but not passing the significance level. In addition, out of 3 DDG associations, one (bisoprolol-atorvastatin-rs9516519 (ABCC4)) was validated.

The lack of replication beyond these findings is to be expected. In part, this is due to the large number of tests and high likelihood to find false positives. This is more likely to be seen with findings from the DDGIs study using the combined Scottish cohort. Sample sizes were generally quite small for most drug combinations to detect reliable statically significance associations. The first reduction in sample sizes was mostly driven by the fact that number of the combination users should be always lower than users of individual drugs. The second larger reduction has occurred

mainly in the case groups where we were looking only for patients who developed rare events (e.g., drug-stop) and within a limited period of time (i.e., the interaction time). The situation was much better with the DGI study as we had relatively larger sample sizes. However, the drug-stop and dose-decrease phenotypes were not very common events so the number of cases quite limited. A greater power and more interpretable results could have been achieved by grouping all individuals with the similar metabolizer status as defined by all variants in the same gene rather than considering each variant individually as in our study. A greater power could have also been achieved by grouping drugs as classes according to their substrate specificities (e.g., all CYP2D6 substrates) rather than examining each drug individually as in our study. However, all of these two methods require a previous knowledge on either the variant function or whether the drug is a known substrate for the gene which counters one of the main targets of this project which is uncovering novel variants and/or novel drug targets.

The second main reason beyond lack of replication is the generic nature of our drug response phenotype that is inherently noisy and non-specific. For example, if the same adverse drug reaction has been observed from the same drug with carriers of the same risky allele, one physician may decide to stop the drug, the second might decide to decrease the dose, the third might keep the drug but provide the patient with a specific lifestyle changes advice to deal with this adverse event, the fourth might keep the drug but prescribe another drug to treat the side effect, and so forth. In short, the effect seen from the initial finding might be 'diluted' in the replication cohort because of this variability in the clinical decisions for the same clinical

scenario due to different patient groups, and clinical practice variation. Figure 19 below simplifies the concept of why these generic phenotypes may not always be replicated.

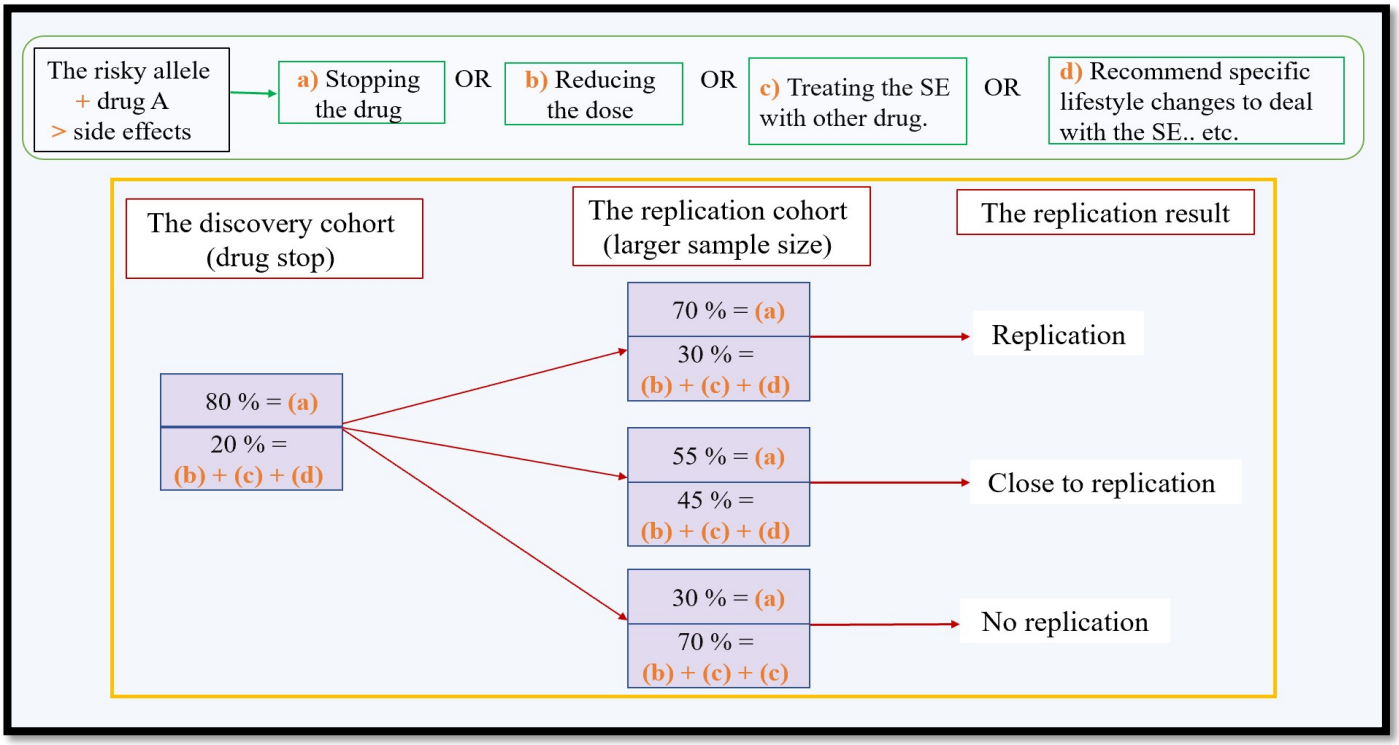


Figure 19: Visual illustration explaining why flexible phenotypes cannot be always replicated.

* Side effects resulting from a certain drug-variant interaction can be dealt with using many strategies (a, b, c, or d). If drug-stop phenotype for example was very common (significant) in the discovery cohort, the same phenotype can be very common (replicated), moderately common (close to replication), or not common (not replicated) in the replication cohort.

On the other hand, this variability is less likely to be seen with drug-specific phenotypes. For example, sulfonylurea-induced hypoglycemia can be defined as blood glucose levels < 4 mmol/l after treatment initiation. This provides accurate identification of all cases of interest as hypoglycemia definition is specifically defined depending on a well-known threshold among health care providers. If a certain variant was associated with this phenotype, then the same association should also be seen in the replication cohort if it was genuine. However, the disadvantage of drug-specific phenotypes is that they are difficult to be utilized to provide a rapid efficient screening of a large variety of DG/DDG associations unlike the 'generic' phenotypes we used in this project. We used these phenotypes as our target was studying a large number of drugs from different drug classes rather than only a specific drug or a specific drug class. The other advantage of 'generic' phenotypes such as 'drug-stop' and 'dose-decrease' is that they can provide a direct evidence for a potential change in the clinical practise, which is the ultimate goal of pharmacogenomics' studies. On the other hand, drug-specific phenotypes may not result in a change in clinical practice. For instance, if a certain variant was found to be associated with increased nicorandil-induced reduction in systolic blood pressure, this provides an evidence for increased risk for hypotension-related ADRs but doesn't necessarily imply that this is also sufficiently clinically significant to result in stopping of or reducing the dose of the treatment.

4. Future directions

Reaching the end of this project is only the beginning of many other exciting scientific journeys. Our findings have identified that ABCC4 transporter as a potential novel transporter for amlodipine and bisoprolol while ABCC3 could be a potential novel transporter for doxazosin. In chapter V, we discussed that quinine was identified as a CYP2C9 substrate even though CYP3A4 was the main metabolizing enzyme. However, we presented another study linking *CYP2C9* genetic variability with the therapeutic efficacy of quinine derivatives chloroquine and primaquine providing more evidence on a potential correlation between *CYP2C9* genetic variants and quinine efficacy/safety as well but further investigation is warranted here. Therefore, it could also be of interest to examine these potential correlations in future work by conducting a laboratory-based investigation. We are collaborating with Kathy Giacomini and colleagues at UCSF to follow up on some of our transporter related hits. They transiently transfect HEK293 Flp-In cells with the transporter and then looks at drug transport into the cell. Regarding metabolizing enzymes-related work, Ronald Wolf's research team, based in Dundee, could help us to identify the degree of correlation between CYP2C9 and quinine metabolism. These in-vitro studies will help us to have a better understanding of the mechanisms by which our findings have occurred. If amlodipine and bisoprolol were ABCC4 substrates, ABCC3 was a transporter for doxazosin, CYP2C9 contributed significantly to quinine metabolism, then our proposed pharmacokinetic mechanisms of interactions I explained in chapter V could be justifiable. However, drug transporters and metabolizing enzymes have many different physiological roles in the body and therefore if these drugs were not substrates for these transporters/enzymes, then there might be other unknown mechanisms of

interactions leading to our observed phenotypes. If we're able to confirm any of the above potential correlations from the laboratory work, then the next step would be to consider conducting genotype-based pharmacokinetic studies. It would be of interest to study the influence of different genotypes at rs4918758 *CYP2C9,* rs9895420 *ABCC3*, and rs868853 *ABCC4* SNPs on quinine, doxazosin, and amlodipine plasma concentrations respectively. If any interesting correlations were detected, then this could be followed by clinical trials in which the outcome to be assessed is dose reduction. Regarding clopidogrel- rs12353214 *PTGS1* SNP association, we could assess the drug-stop phenotype, or bleeding time, or use population data to look at the effect on stroke recurrence, ACS or bleeding risk.

A clinical trial to establish a clinically relevant drug-gene interaction could include a number of different study designs including, for example, an exposure-only design (EOD) [316] or an enrichment design (ED) [317]. The first study design focuses on the treatment group only. In our case, patients are treated with the drug of interest first (i.e., quinine, doxazosin, amlodipine, or clopidogrel), and then these patients are followed up. I then evaluate the outcome of interest (dose-decrease for quinine, doxazosin, or amlodipine and drug-stop for clopidogrel) when the genotype group is known ((rs4918758 (quinine), rs9895420 (doxazosin), rs868853 (amlodipine), or rs12353214 (clopidogrel)).

The other design (ED) starts by knowing the patient genotypes and defining patient subgroups based upon carriage of the variant allele classified into carriers or non-carriers of the variant allele. Then, both groups are treated with the drug of interest and followed up to evaluate the phenotype of interest (dose-decrease/drug-stop) in each group. A significant difference in the number of events between the two groups could indicate a significant drug-variant interaction.

Regarding the bisoprolol-atorvastatin-rs9516519 *ABCC4* SNP-drug-stop association, the drug-stop phenotype could be too rare and tricky to study clinically in DDGIs studies. However, I have validated this finding using an alternative phenotype (SBP drop) which can be investigated better clinically. Here, patients are stratified into two groups: carriers and non-carriers of the variant allele. Then, each group is treated (in random order) with bisoprolol alone and atorvastatin + bisoprolol. Then, I study whether there is any significant difference in SBP levels between the two groups during the bisoprolol-only period (DGI) and during atorvastatin + bisoprolol period (DDGI).

The UKBB primary care data is a very rich resource combining both the large cohort size and the large variety of clinical phenotypes such as clinical measurements, direct drug-related adverse events, and indirect but could be drug-related adverse events. This data only became available towards the end of my PhD and is likely to soon incorporate the primary care data for all 500,000 participants making it the world's largest longitudinal data for pharmacogenetic studies. This will facilitate studying a large variety of drug response phenotypes in a large cohort leading to more previously undiscovered associations. In the future, I would like to use this database to study "drug-stop" and "dose-decrease" phenotypes again for all of my selected drugs and drug combinations. I would also like to take a more targeted drug-specific approach to look at drug-specific adverse events or drug efficacy using clinical measurements such as blood pressure change, HbA1c change or change in electrolytes.

Scientific research will be an important part of my academic responsibilities in the university where I will work in Saudi Arabia. Therefore, UKBB primary care data offers considerable potential to support me in this work. Another interesting database

I propose to work with includes patients from Saudi Arabia - Genetic and pharmacogenomic studies are ultra-rare in the Arabic region making studies on this particular ethnic group of great importance and of course relevant to the management of patients in Saudi Arabia. The Saudi Human Genome Program (SHGP) [318] is the largest genetic project in the middle east area aiming to collect genotype data for 100,000 Saudi individuals to facilitate the discovery of genetic variants associated with rare diseases in Saudi individuals in addition to pharmacogenomic studies. The project started in 2013, and the number of participants to date (2020) reached 56,799 individuals [319]. My future plan is to establish a collaborative research work between the incredible research team I worked with in the University of Dundee and the department of Clinical Pharmacy in the College of Pharmacy in Umm Al-Qura University where I will work to facilitate pharmacogenomic research on both the British and Saudi populations.

Publications and prizes

1. I have been awarded the **1st prize award** (*see supplementary material 3*) for the best poster presentation and research outcome among all other submissions for my work titled "*Drug-Gene and Drug-Drug-Gene Interactions for the Most Commonly Used Drugs in the UK"* during the Postgraduate Research Student Symposium 2019 which has been held in the University of Dundee. I have received an **encouragement award of ~ £ 1000** from the Royal Embassy of Saudi Arabia in London upon this achievement.
2. One review article has been **published** at the end of 2019.

   " **Malki M**, Pearson E. Drug–drug–gene interactions and adverse drug reactions. *The Pharmacogenomics Journal*. 2019;20(3):355-366 ". The paper has been **highlighted by the journal as one of the best publications** in *The Pharmacogenomic Journal 2019* .I have received an **encouragement award of ~ £ 1000** from the Royal Embassy of Saudi Arabia in London upon this paper publication.

   Regarding publishing our main findings from this project, we preferred not to publish our work until after we can support it with a replication study from an independent cohort. The replication cohort (the UKBB primary care data) wasn't available to us until after the end of my 3rd year in this 4-years PhD program. The last (4th year) started with working with this data, followed by the writing-up stage of my thesis. Therefore, I couldn't find enough time to publish our main findings. However, two papers, one for DGIs and the other one for DDGIs for the most commonly used drugs in the UK, have already been written. We plan to publish the DGI paper in the near future.

References

Chapter I

1. Lazarou J, Pomernaz B, Corey P. Incidence of Adverse Drug Reactions in Hospitalized Patients: A Meta-Analysis of Prospective Studies. Survey of Anesthesiology. 1999;43(1):53-54.
2. Bouvy J, De Bruin M, Koopmanschap M. Epidemiology of Adverse Drug Reactions in Europe: A Review of Recent Observational Studies. Drug Safety. 2015;38(5):437-453.
3. MiDatabank Adverse Drug Reaction (ADR) reporting – 2016 update – SPS - Specialist Pharmacy Service – The first stop for professional medicines advice [Internet]. Sps.nhs.uk. 2017 [cited 6 September 2017]. Available from: https://www.sps.nhs.uk/articles/midatabank-adverse-drug-reaction-adr-reporting-2016-update/
4. Exploring the costs of unsafe care in the NHS [Internet]. frontier-economics. 2014 [cited 11 September 2017]. Available from: http://www.frontier-economics.com/documents/2014/10/exploring-the-costs-of-unsafe-care-in-the-nhs-frontier-report-2-2-2-2.pdf
5. Howard R, Avery A, Slavenburg S, Royal S, Pipe G, Lucassen P et al. Which drugs cause preventable admissions to hospital? A systematic review. British Journal of Clinical Pharmacology. 2007;63(2):136-147.
6. Rodrigues M, Oliveira C. Drug-drug interactions and adverse drug reactions in polypharmacy among older adults: an integrative review. Revista Latino-Americana de Enfermagem. 2016; 24(0):e2800.
7. Garrod A. ESP Digital Books: Inborn Errors of Metabolism [Internet]. Esp.org. 1923 [cited 26 September 2020]. Available from: http://www.esp.org/books/garrod/inborn-errors/facsimile/
8. Lehmann H, Ryan E. THE FAMILIAL INCIDENCE OF LOW PSEUDOCHOLINESTERASE LEVEL. The Lancet. 1956;268(6944):681-682.
9. Motulsky A. DRUG REACTIONS, ENZYMES, AND BIOCHEMICAL GENETICS. Journal of the American Medical Association. 1957;165(7):835.
10. Vogel F. Moderne Probleme der Humangenetik. Ergebnisse der Inneren Medizin und Kinderheilkunde. 1959;12:52-125.
11. Verbeurgt P, Mamiya T, Oesterheld J. How common are drug and gene interactions? Prevalence in a sample of 1143 patients with *CYP2C9*, *CYP2C19* and *CYP2D6* genotyping. Pharmacogenomics. 2014;15(5):655-665.
12. Bahar M, Setiawan D, Hak E, Wilffert B. Pharmacogenetics of drug–drug interaction and drug–drug–gene interaction: a systematic review on *CYP2C9*, *CYP2C19* and *CYP2D6*. Pharmacogenomics. 2017;18(7):701-739.

13. Wizemann T, Berger A. The value of genetic and genomic technologies. Washington, D.C.: National Academies Press; 2010.
14. FDA. Drugs@FDA: FDA-Approved Drugs [Internet]. Accessdata.fda.gov. 2020 [cited 26 September 2020]. Available from: https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm?event=overview.process&varApplNo=009218
15. Shah R. Genotype-guided warfarin therapy: Still of only questionable value two decades on. Journal of Clinical Pharmacy and Therapeutics. 2020;45(3):547-560.
16. Pirmohamed M, Burnside G, Eriksson N, Jorgensen A, Toh C, Nicholson T et al. A Randomized Trial of Genotype-Guided Dosing of Warfarin. New England Journal of Medicine. 2013;369(24):2294-2303.
17. Goetz M, Sangkuhl K, Guchelaar H, Schwab M, Province M, Whirl-Carrillo M, Symmans W, McLeod H, Ratain M, Zembutsu H, Gaedigk A, van Schaik R, Ingle J, Caudle K, Klein T. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for *CYP2D6* and Tamoxifen Therapy. Clinical Pharmacology & Therapeutics. 2018;103(5):770-777.
18. Lima J, Thomas C, Barbarino J, Desta Z, Van Driest S, El Rouby N, Johnson J, Cavallari L, Shakhnovich V, Thacker D, Scott S, Schwab M, Uppugunduri C, Formea C, Franciosi J, Sangkuhl K, Gaedigk A, Klein T, Gammal R, Furuta T. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for *CYP2C19* and Proton Pump Inhibitor Dosing. Clinical Pharmacology & Therapeutics. 2020.
19. Crews K, Gaedigk A, Dunnenberger H, Leeder J, Klein T, Caudle K, Haidar C, Shen D, Callaghan J, Sadhasivam S, Prows C, Kharasch E, Skaar T. Clinical Pharmacogenetics Implementation Consortium Guidelines for Cytochrome P450 2D6 Genotype and Codeine Therapy: 2014 Update. Clinical Pharmacology & Therapeutics. 2014;95(4):376-382.
20. Andersson M, Eliasson E, Lindh J. A clinically significant interaction between warfarin and simvastatin is unique to carriers of the*CYP2C9**3allele. Pharmacogenomics. 2012;13(7):757-762.
21. Malhi H. Warfarin and celecoxib interaction in the setting of cytochrome P450 (*CYP2C9*) polymorphism with bleeding complication. Postgraduate Medical Journal. 2004;80(940):107-109.
22. Uno T, Shimizu M, Yasui-Furukori N, Sugawara K, Tateishi T. Different effects of fluvoxamine on rabeprazole pharmacokinetics in relation to *CYP2C19* genotype status. British Journal of Clinical Pharmacology. 2006;61(3):309-314.
23. Yasui-Furukori N, Saito M, Uno T, Takahata T, Sugawara K, Tateishi T. Effects of Fluvoxamine on Lansoprazole Pharmacokinetics in Relation to *CYP2C19* Genotypes. The Journal of Clinical Pharmacology. 2004;44(11):1223-1229.

24. Cho J, Yu K, Jang I, Yang B, Shin S, Yim D. Omeprazole hydroxylation is inhibited by a single dose of moclobemide in homozygotic EM genotype for *CYP2C19*. British Journal of Clinical Pharmacology. 2002;53(4):393-397.
25. Hamelin B. Significant interaction between the nonprescription antihistamine diphenhydramine and the *CYP2D6* substrate metoprolol in healthy men with high or low *CYP2D6* activity. Clinical Pharmacology & Therapeutics. 2000;67(5):466-477.
26. Damy T, Pousset F, Caplain H, Hulot J, Lechat P. Pharmacokinetic and pharmacodynamic interactions between metoprolol and dronedarone in extensive and poor *CYP2D6* metabolizers healthy subjects. Fundamental and Clinical Pharmacology. 2004;18(1):113-123.
27. Zhu L, Brüggemann R, Uy J, Colbers A, Hruska M, Chung E et al. *CYP2C19* Genotype-Dependent Pharmacokinetic Drug Interaction Between Voriconazole and Ritonavir-Boosted Atazanavir in Healthy Subjects. J Clin Pharmacol. 2017;57(2):235-246.
28. Imamura C, Furihata K, Okamoto S, Tanigawara Y. Impact of cytochrome P450 2C19 polymorphisms on the pharmacokinetics of tacrolimus when coadministered with voriconazole. The Journal of Clinical Pharmacology. 2015;56(4):408-413.
29. 12. Shi H, Yan J, Zhu W, Yang G, Tan Z, Wu W et al. Effects of erythromycin on voriconazole pharmacokinetics and association with *CYP2C19* polymorphism. European Journal of Clinical Pharmacology. 2010;66(11):1131-1136.
30. 13. Saito M, Yasui-Furukori N, Uno T, Takahata T, Sugawara K, Munakata A et al. Effects of clarithromycin on lansoprazole pharmacokinetics between *CYP2C19* genotypes. British Journal of Clinical Pharmacology. 2005;59(3):302-309.
31. Sangkuhl K, Klein T, Altman R. Clopidogrel pathway. Pharmacogenetics and Genomics. 2010; 20(7):463-465.
32. Furuta T, Iwaki T, Umemura K. Influences of different proton pump inhibitors on the anti-platelet function of clopidogrel in relation to *CYP2C19* genotypes. British Journal of Clinical Pharmacology. 2010;70(3):383-392.
33. Harmze A, Van Werkum J, Souverein P, Breet N, Bouman H, Hackeng C et al. Combined influence of proton-pump inhibitors, calcium-channel blockers and *CYP2C19**2 on on-treatment platelet reactivity and on the occurrence of atherothrombotic events after percutaneous coronary intervention. Journal of Thrombosis and Haemostasis. 2011;9(10):1892-1901.
34. Malingré M, Godschalk P, Klein S. A case report of voriconazole therapy failure in a homozygous ultrarapid *CYP2C19**17/*17 patient comedicated with carbamazepine. British Journal of Clinical Pharmacology. 2012;74(1):205-206.

35. Carlquist J, Knight S, Horne B, Huntinghouse J, Rollo J, Muhlestein J et al. Cardiovascular risk among patients on clopidogrel anti-platelet therapy after placement of drug-eluting stents is modified by genetic variants in both the *CYP2C19* and *ABCB1* genes. Thrombosis and Haemostasis. 2013;109(4):744-754.
36. Chan M, Tan K, Tan H, Huan P, Li B, Phua Q et al. *CYP2C19* and PON1 polymorphisms regulating clopidogrel bioactivation in Chinese, Malay and Indian subjects. Pharmacogenomics. 2012;13(5):533-542.
37. Harmsze A, van Werkum J, Hackeng C, Ruven H, Kelder J, Bouman H et al. The influence of *CYP2C19**2 and *17 on on-treatment platelet reactivity and bleeding events in patients undergoing elective coronary stenting. Pharmacogenetics and Genomics. 2012;22(3):169-175.
38. Cayla G, Hulot J, O'Connor S, Pathak A, Scott S, Gruel Y et al. Clinical, angiographic, and genetic factors associated with early coronary stent thrombosis. JAMA. 2011;306(16):1765-74.
39. Gurbel P, Shuldiner A, Bliden K, Ryan K, Pakyz R, Tantry U. The relation between *CYP2C19* genotype and phenotype in stented patients on maintenance dual antiplatelet therapy. American Heart Journal. 2011;161(3):598-604.
40. Wallentin L, James S, Storey R, Armstrong M, Barratt B, Horrow J et al. Effect of *CYP2C19* and *ABCB1* single nucleotide polymorphisms on outcomes of treatment with ticagrelor versus clopidogrel for acute coronary syndromes: a genetic substudy of the PLATO trial. The Lancet. 2010;376(9749):1320-1328.
41. Sibbing D, Gebhard D, Koch W, Braun S, Stegherr J, Morath T et al. Isolated and interactive impact of common *CYP2C19* genetic variants on the antiplatelet effect of chronic clopidogrel therapy. Journal of Thrombosis and Haemostasis. 2010;8(8):1685-1693.
42. Gladding P, Webster M, Zeng I, Farrell H, Stewart J, Ruygrok P et al. The Pharmacogenomics and Pharmacodynamics of Clopidogrel Response: An Analysis From the PRINC (Plavix Response in Coronary Intervention) Trial. Heart, Lung and Circulation. 2008;17:S82.
43. Dai Z. Relationship between cytochrome P450 2C19*17 genotype distribution platelet aggregation and bleeding risk in patients with blood stasis syndrome of coronary artery disease treated with clopidogrel. Journal of Chinese Integrative Medicine. 2012;10(6):647-654.
44. Tiroch K, Sibbing D, Koch W, Roosen-Runge T, Mehilli J, Schömig A et al. Protective effect of the *CYP2C19* *17 polymorphism with increased activation of clopidogrel on cardiovascular events. American Heart Journal. 2010;160(3):506-512.
45. Vormfelde S, Brockmöller J, Bauer S, Herchenhein P, Kuon J, Meineke I et al. Relative Impact of Genotype and Enzyme Induction on the Metabolic Capacity of *CYP2C9* in Healthy Volunteers. Clinical Pharmacology & Therapeutics. 2009;86(1):54-61.

46. Depta J, Lenzini P, Lanfear D, Wang T, Spertus J, Bach R et al. Clinical outcomes associated with proton pump inhibitor use among clopidogrel-treated patients within *CYP2C19* genotype groups following acute myocardial infarction. The Pharmacogenomics Journal. 2014;15(1):20-25.
47. Laine K, Tybring G, Hartter S, Andersson K, Svensson J, Widen J et al. Inhibition of cytochrome P4502D6 activity with paroxetine normalizes the ultrarapid metabolizer phenotype as measured by nortriptyline pharmacokinetics and the debrisoquin test. Clinical Pharmacology & Therapeutics. 2001;70(4):327-335.
48. Nicholls G, Youdim K. Drug Transporters: Volume 1: Role and Importance in ADME and Drug Development, Figure 1.3, p.16. [Internet]. 54th ed. The Royal Society of Chemistry (RSC) 2016; 2016 [cited 9 September 2017]. Available from: http://pubs.rsc.org/en/content/ebook/978-1-78262-069-3
49. PHARMGKB. rs1045642 [Internet]. PharmGKB. 2020 [cited 26 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166157284/clinicalAnnotation/1296599132
50. PHARMGKB. rs1045642 [Internet]. PharmGKB. 2020 [cited 26 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166157284/clinicalAnnotation/655386244
51. PHARMGKB. rs1045642 [Internet]. PharmGKB. 2020 [cited 26 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166157284/clinicalAnnotation
52. PHARMGKB. rs2231142 [Internet]. PharmGKB. 2020 [cited 26 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166156544/clinicalAnnotation/1154221922
53. Wessler J, Grip L, Mendell J, Giugliano R. The P-Glycoprotein Transport System and Cardiovascular Drugs. Journal of the American College of Cardiology. 2013;61(25):2495-2502.
54. Wang Y, Li J, Wang X, Zhang Y, Wang C, Huang M. Diltiazem augments the influence of MDR1 genotype status on cyclosporine concentration in Chinese patients with renal transplantation. Acta Pharmacologica Sinica. 2015;36(7):855-862.
55. Uehlinger C, Crettol S, Chassot P, Brocard M, Koeb L, Brawand-Amey M et al. Increased (R)-Methadone Plasma Concentrations by Quetiapine in Cytochrome P450s and *ABCB1* Genotyped Patients. Journal of Clinical Psychopharmacology. 2007;27(3):273-278.
56. Tsuji D, Kim Y, Nakamichi H, Daimon T, Suwa K, Iwabe Y et al. Association of *ABCB1* Polymorphisms with the Antiemetic Efficacy of Granisetron plus Dexamethasone in Breast Cancer Patients. Drug Metabolism and Pharmacokinetics. 2013;28(4):299-304.
57. Gouin-Thibault I, Delavenne X, Blanchard A, Siguret V, Salem J, Narjoz C et al. Interindividual variability in dabigatran and rivaroxaban exposure: contribution of*ABCB1*genetic polymorphisms and interaction with clarithromycin. Journal of Thrombosis and Haemostasis. 2017;15(2):273-283.

58. Togashi M, Niioka T, Komatsuda A, Nara M, Okuyama S, Omokawa A et al. Effect of *CYP3A5* and *ABCB1* polymorphisms on the interaction between tacrolimus and itraconazole in patients with connective tissue disease. European Journal of Clinical Pharmacology. 2015;71(9):1091-1097.
59. Lowenberg D, Thorn C, Whirl-Carrillo M, Ramirez J, Gong L, Marsh S et al. Irinotecan Pathway, Pharmacokinetics [Internet]. PharmGKB. [cited 9 October 2017]. Available from: https://www.pharmgkb.org/pathway/PA2001
60. Li M, Seiser E, Baldwin R, Ramirez J, Ratain M, Innocenti F et al. ABC transporter polymorphisms are associated with irinotecan pharmacokinetics and neutropenia. The Pharmacogenomics Journal. 2016. doi: 10.1038/tpj.2016.75.
61. Warren R, Smith R, Campalani E, Eyre S, Smith C, Barker J et al. Genetic Variation in Efflux Transporters Influences Outcome to Methotrexate Therapy in Patients with Psoriasis. Journal of Investigative Dermatology. 2008;128(8):1925-1929.
62. Jungsuwadee P, Nithipongvanitch R, Chen Y, Oberley T, Butterfield D, St. Clair D et al. Mrp1 Localization and Function in Cardiac Mitochondria after Doxorubicin. Molecular Pharmacology. 2009;75(5):1117-1126.
63. Wojnowski L. NAD(P)H Oxidase and Multidrug Resistance Protein Genetic Polymorphisms Are Associated With Doxorubicin-Induced Cardiotoxicity. Circulation. 2005;112(24):3754-3762.
64. Kobayashi K, Ito K, Takada T, Sugiyama Y, Suzuki H. Functional analysis of nonsynonymous single nucleotide polymorphism type ATP-binding cassette transmembrane transporter subfamily C member 3. Pharmacogenetics and Genomics. 2008;18(9):823-833.
65. Choi H, Bae K, Cho S, Ghim J, Choe S, Jung J et al. Impact of *CYP2D6*, *CYP3A5*, *CYP2C19*, *CYP2A6*, *SLCO1B1*, *ABCB1*, and *ABCG2* gene polymorphisms on the pharmacokinetics of simvastatin and simvastatin acid. Pharmacogenet Genomics. 2013;25(12):595-608.
66. Luzum J, Theusch E, Taylor K, Wang A, Sadee W, Binkley P et al. Individual and Combined Associations of Genetic Variants in *CYP3A4*, *CYP3A5*, and *SLCO1B1* With Simvastatin and Simvastatin Acid Plasma Concentrations. Journal of Cardiovascular Pharmacology. 2015;66(1):80-85.
67. Mirošević Skvrce N, Macolić Šarinić V, Šimić I, Ganoci L, Muačević Katanec D, Božina N. *ABCG2* gene polymorphisms as risk factors for atorvastatin adverse reactions: a case-control study. Pharmacogenomics. 2015;16(8):803-815.
68. Tsamandouras N, Dickinson G, Guo Y, Hall S, Rostami-Hodjegan A, Galetin A et al. Development and Application of a Mechanistic Pharmacokinetic Model for Simvastatin and its Active Metabolite Simvastatin Acid Using an Integrated Population PBPK Approach. Pharmaceutical Research. 2014;32(6):1864-1883.

69. Birmingham B, Bujac S, Elsby R, Azumaya C, Wei C, Chen Y et al. Impact of *ABCG2* and *SLCO1B1* polymorphisms on pharmacokinetics of rosuvastatin, atorvastatin and simvastatin acid in Caucasian and Asian subjects: a class effect?. European Journal of Clinical Pharmacology. 2015;71(3):341-355.
70. Tsamandouras N, Dickinson G, Guo Y, Hall S, Rostami-Hodjegan A, Galetin A et al. Identification of the Effect of Multiple Polymorphisms on the Pharmacokinetics of Simvastatin and Simvastatin Acid Using a Population-Modeling Approach. Clinical Pharmacology & Therapeutics. 2014;96(1):90-100.
71. Ferrari M, Guasti L, Maresca A, Mirabile M, Contini S, Grandi A et al. Association between statin-induced creatine kinase elevation and genetic polymorphisms in *SLCO1B1*, *ABCB1* and *ABCG2*. European Journal of Clinical Pharmacology. 2014;70(5):539-547.
72. De Keyser C, Peters B, Becker M, Visser L, Uitterlinden A, Klungel O et al. The *SLCO1B1* c.521T>C polymorphism is associated with dose-decrease or switching during statin therapy in the Rotterdam Study. Pharmacogenetics and Genomics. 2014;24(1):43-51.
73. DeGorter M, Tirona R, Schwarz U, Choi Y, Dresser G, Suskin N et al. Clinical and Pharmacogenetic Predictors of Circulating Atorvastatin and Rosuvastatin Concentrations in Routine Clinical Care. Circulation: Cardiovascular Genetics. 2013;6(4):400-408.
74. Carr D, O'Meara H, Jorgensen A, Campbell J, Hobbs M, McCann G et al. *SLCO1B1* Genetic Variant Associated With Statin-Induced Myopathy: A Proof-of-Concept Study Using the Clinical Practice Research Datalink. Clinical Pharmacology & Therapeutics. 2013;94(6):695-701.
75. Ulvestad M, Skottheim I, Jakobsen G, Bremer S, Molden E, Åsberg A et al. Impact of OATP1B1, MDR1, and *CYP3A4* Expression in Liver and Intestine on Interpatient Pharmacokinetic Variability of Atorvastatin in Obese Subjects. Clinical Pharmacology & Therapeutics. 2012;93(3):275-282.
76. Marciante K, Durda J, Heckbert S, Lumley T, Rice K, McKnight B et al. Cerivastatin, genetic variants, and the risk of rhabdomyolysis. Pharmacogenet Genomics. 2011;21(5):280-288.
77. Donnelly L, Doney A, Tavendale R, Lang C, Pearson E, Colhoun H et al. Common Nonsynonymous Substitutions in *SLCO1B1* Predispose to Statin Intolerance in Routinely Treated Individuals With Type 2 Diabetes: A Go-DARTS Study. Clinical Pharmacology & Therapeutics. 2010;89(2):210-216.
78. Voora D, Shah S, Spasojevic I, Ali S, Reed C, Salisbury B et al. The *SLCO1B1**5 genetic variant is associated with statin-induced side effects. J Am Coll Cardiol. 2009;54(17):1609-1616.
79. Link E, Parish S, Armitage J, Bowman L, Heath S, Matsuda F et al. *SLCO1B1* variants and statin-induced myopathy--a genomewide study. N Engl J Med. 2008;359(8):789-799.

80. Ho R, Choi L, Lee W, Mayo G, Schwarz U, Tirona R et al. Effect of drug transporter genotypes on pravastatin disposition in European- and African-American participants. Pharmacogenetics and Genomics. 2007;17(8):647-656.
81. Pasanen M, Fredrikson H, Neuvonen P, Niemi M. Different Effects of *SLCO1B1* Polymorphism on the Pharmacokinetics of Atorvastatin and Rosuvastatin. Clinical Pharmacology & Therapeutics. 2007;82(6):726-733.
82. Niemi M, Pasanen M, Neuvonen P. *SLCO1B1* polymorphism and sex affect the pharmacokinetics of pravastatin but not fluvastatin. Clinical Pharmacology & Therapeutics. 2006;80(4):356-366.
83. Igel M, Arnold K, Niemi M, Hofmann U, Schw AB M, Lutjohann D et al. Impact of the *SLCO1B1* polymorphism on the pharmacokinetics and lipid-lowering efficacy of multiple-dose pravastatin. Clinical Pharmacology & Therapeutics. 2006;79(5):419-426.
84. Pasanen M, Neuvonen M, Neuvonen P, Niemi M. *SLCO1B1* polymorphism markedly affects the pharmacokinetics of simvastatin acid. Pharmacogenetics and Genomics. 2006;16(12):873-879.
85. Thompson J, Man M, Johnson K, Wood L, Lira M, Lloyd D et al. An association study of 43 SNPs in 16 candidate genes with atorvastatin response. The Pharmacogenomics Journal. 2005;5(6):352-358.
86. Mwinyi J. Evidence for inverse effects of OATP-C (SLC21A6) *5 and *1b haplotypes on pravastatin kinetics. Clinical Pharmacology & Therapeutics. 2004;75(5):415-421.
87. Nishizato Y. Polymorphisms of OATP-C (SLC21A6) and OAT3 (*SLC22A8*) genes: Consequences for pravastatin pharmacokinetics. Clinical Pharmacology & Therapeutics. 2003;73(6):554-565.
88. Ramsey L, Johnson S, Caudle K, Haidar C, Voora D, Wilke R, Maxwell W, McLeod H, Krauss R, Roden D, Feng Q, Cooper-DeHoff R, Gong L, Klein T, Wadelius M, Niemi M. The Clinical Pharmacogenetics Implementation Consortium Guideline for *SLCO1B1* and Simvastatin-Induced Myopathy: 2014 Update. Clinical Pharmacology & Therapeutics. 2014;96(4):423-428.
89. Dujic T, Zhou K, Donnelly L, Tavendale R, Palmer C, Pearson E. Association of Organic Cation Transporter 1 With Intolerance to Metformin in Type 2 Diabetes: A GoDARTS Study. Diabetes. 2014;64(5):1786-1793.
90. Ahlin G, Chen L, Lazorova L, Chen Y, Ianculescu A, Davis R et al. Genotype-dependent effects of inhibitors of the organic cation transporter, OCT1: predictions of metformin interactions. The Pharmacogenomics Journal. 2010;11(6):400-411.
91. Grün B, Kiessling M, Burhenne J, Riedel K, Weiss J, Rauch G et al. Trimethoprim-metformin interaction and its genetic modulation by OCT2 and MATE1 transporters. British Journal of Clinical Pharmacology. 2013;76(5):787-796.

92. Wang Z, Yin O, Tomlinson B, Chow M. OCT2 polymorphisms and in-vivo renal functional consequence: studies with metformin and cimetidine. Pharmacogenetics and Genomics. 2008;18(7):637-645.
93. Filipski K, Mathijssen R, Mikkelsen T, Schinkel A, Sparreboom A. Contribution of Organic Cation Transporter 2 (OCT2) to Cisplatin-Induced Nephrotoxicity. Clinical Pharmacology & Therapeutics. 2009;86(4):396-402.
94. Spracklen T, Vorster A, Ramma L, Dalvie S, Ramesar R. Promoter region variation in NFE2L2 influences susceptibility to ototoxicity in patients exposed to high cumulative doses of cisplatin. The Pharmacogenomics Journal. 2016; doi: 10.1038/tpj.2016.52.
95. Lanvers-Kaminsky C, Sprowl J, Malath I, Deuster D, Eveslage M, Schlatter E et al. Human OCT2 variant c.808G>T confers protection effect against cisplatin-induced ototoxicity. Pharmacogenomics. 2015;16(4):323-332.
96. Aquilante C, Kiser J, Anderson P, Christians U, Kosmiski L, Daily E et al. Influence of *SLCO1B1*Polymorphisms on the Drug-Drug Interaction Between Darunavir/Ritonavir and Pravastatin. The Journal of Clinical Pharmacology. 2012;52(11):1725-1738.
97. Suwannakul S, Ieiri I, Kimura M, Kawabata K, Kusuhara H, Hirota T et al. Pharmacokinetic interaction between pravastatin and olmesartan in relation to *SLCO1B1* polymorphism. Journal of Human Genetics. 2008;53(10):899-904.
98. Kalliokoski A, Backman J, Kurkinen K, Neuvonen P, Niemi M. Effects of Gemfibrozil and Atorvastatin on the Pharmacokinetics of Repaglinide in Relation to *SLCO1B1* Polymorphism. Clinical Pharmacology & Therapeutics. 2008;84(4):488-496.
99. He Y, Zhang W, Chen Y, Guo D, Tu J, Xu L et al. Rifampicin alters atorvastatin plasma concentration on the basis of *SLCO1B1* 521T>C polymorphism. Clinica Chimica Acta. 2009;405(1-2):49-52.
100. Cho S, Chung J. The MATE1 rs2289669 polymorphism affects the renal clearance of metformin following ranitidine treatment. Int Journal of Clinical Pharmacology and Therapeutics. 2016;54(04):253-262.
101. Zhou K, Donnelly L, Burch L, Tavendale R, Doney A, Leese G et al. Loss-of-Function *CYP2C9* Variants Improve Therapeutic Response to Sulfonylureas in Type 2 Diabetes: A Go-DARTS Study. Clinical Pharmacology & Therapeutics. 2009;87(1):52-56.
102. Ragia G, Petridis I, Tavridou A, Christakidis D, Manolopoulos V. Presence of *CYP2C9**3 allele increases risk for hypoglycemia in Type 2 diabetic patients treated with sulfonylureas. Pharmacogenomics. 2009;10(11):1781-1787.
103. Jiang F, Choi J, Lee J, Ryu S, Park Z, Lee J et al. The influences of *SLCO1B1* and *ABCB1* genotypes on the pharmacokinetics of simvastatin, in relation to *CYP3A4* inhibition. Pharmacogenomics. 2017;18(5):459-469.

104. Meyer zu Schwabedissen H, Siegmund W, Kroemer H, Rollnik J. Creatine kinase elevation caused by a combination of fluvastatin and telmisartan in a patient heterozygous for the *CYP2C9**3 and *ABCC2* -24C > T variants: a case report. BMC Research Notes. 2014;7(1):688.

105. Marusic S, Lisicic A, Horvatic I, Bacic-Vrca V, Bozina N. Atorvastatin-related rhabdomyolysis and acute renal failure in a genetically predisposed patient with potential drug–drug interaction. International Journal of Clinical Pharmacy. 2012;34(6):825-827.

106. Rosuvastatin - DrugBank [Internet]. Drugbank.ca. [cited 9 October 2017]. Available from: https://www.drugbank.ca/drugs/DB01098

107. Tod M, Nkoud-Mongo C, Gueyffier F. Impact of Genetic Polymorphism on Drug-Drug Interactions Mediated by Cytochromes: A General Approach. The AAPS Journal. 2013;15(4):1242-1252.

108. Zhuang X, Lu C. PBPK modeling and simulation in drug research and development. Acta Pharmaceutica Sinica B. 2016;6(5):430-440.

109. Mariappan T, Shen H, Marathe P. Endogenous Biomarkers to Assess Drug-Drug Interactions by Drug Transporters and Enzymes. Current Drug Metabolism. 2017; 18(8):757-768.

110. Magro L, Moretti U, Leone R. Epidemiology and characteristics of adverse drug reactions caused by drug–drug interactions. Expert Opinion on Drug Safety. 2011;11(1):83-94

Chapter II

111. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J et al. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. PLOS Medicine. 2015;12(3):e1001779.

112. Bycroft C, Freeman C, Petkova D, Band G, T. Elliott L, Sharp K et al. Genome-wide genetic data on ~500,000 UK Biobank participants. biorxiv [Internet]. 2017 [cited 26 September 2020];. Available from: https://www.biorxiv.org/content/10.1101/166298v1

113. Hébert H, Shepherd B, Milburn K, Veluchamy A, Meng W, Carr F et al. Cohort Profile: Genetics of Diabetes Audit and Research in Tayside Scotland (GoDARTS). International Journal of Epidemiology. 2017;47(2):380-381j.

114. Smith B, Campbell A, Linksted P, Fitzpatrick B, Jackson C, Kerr S et al. Cohort Profile: Generation Scotland: Scottish Family Health Study (GS:SFHS). The study, its participants and their potential for genetic research on health and illness. International Journal of Epidemiology. 2012;42(3):689-700.

115. Kerr S, Campbell A, Murphy L, Hayward C, Jackson C, Wain L et al. Pedigree and genotyping quality analyses of over 10,000 DNA samples from the Generation Scotland: Scottish Family Health Study. BMC Medical Genetics. 2013;14(1).

116. The University of Edinburgh. The Scottish Family Health Study [Internet]. The University of Edinburgh. 2020 [cited 26 September 2020]. Available from: https://www.ed.ac.uk/generation-scotland/for-researchers/generation-scotland-studies/scottish-family-health-study

117. McKinstry B, Sullivan F, Vasishta S, Armstrong R, Hanley J, Haughney J et al. Cohort profile: the Scottish Research register SHARE. A register of people interested in research participation linked to NHS data sets. BMJ Open. 2017;7(2):e013351.

118. SHARE. Share | Register4Share [Internet]. Registerforshare.org. 2020 [cited 26 September 2020]. Available from: https://www.registerforshare.org/

119. University of Dundee. GoSHARE is after your blood - help world-leading research in Tayside [Internet]. Discovery - the University of Dundee Research Portal. 2020 [cited 26 September 2020]. Available from: https://discovery.dundee.ac.uk/en/clippings/goshare-is-after-your-blood-help-world-leading-research-in-taysid

120. Padmanabhan S. Handbook of pharmacogenomics and stratified medicine. 1st ed. Glasgow ,UK: Elsevier Academic Press; 2017.

121. Winter B. Linear models and linear mixed effect models in R with linguistic applications [Internet]. 2013 [cited 27 September 2020]. Available from: http://www.bodowinter.com/uploads/1/2/9/3/129362560/bw_lme_tutorial1.pdf

122. Winter B. Linear models and linear mixed effects models in R with linguistic applications [Internet]. 2013 [cited 27 September 2020]. Available from: http://www.bodowinter.com/uploads/1/2/9/3/129362560/bw_lme_tutorial2.pdf

123. R L. Logistic Regression Assumptions and Diagnostics in R - Articles - STHDA [Internet]. Sthda.com. 2018 [cited 27 September 2020]. Available from: http://www.sthda.com/english/articles/36-classification-methods-essentials/148-logistic-regression-assumptions-and-diagnostics-in-r/

124. Wishart D, Feunang Y, Guo A, Lo E, Marcu A, Grant J et al. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Research. 2017;46(D1):D1074-D1082.

125. Barbarino J, Whirl-Carrillo M, Altman R, Klein T. PharmGKB: A worldwide resource for pharmacogenomic information. Wiley Interdisciplinary Reviews: Systems Biology and Medicine. 2018;10(4):e1417.

126. Whirl-Carrillo M, McDonagh E, Hebert J, Gong L, Sangkuhl K, Thorn C et al. Pharmacogenomics Knowledge for Personalized Medicine. Clinical Pharmacology & Therapeutics. 2012;92(4):414-417.

127. Canela-Xandri O, Rawlik K, Tenesa A. An atlas of genetic associations in UK Biobank. Nature Genetics. 2018;50(11):1593-1599.

Chapter III

128. PHARMGKB. VIPs: Very Important Pharmacogenes [Internet]. PharmGKB. 2020 [cited 27 September 2020]. Available from: https://www.pharmgkb.org/vips

129. Ragia G, Petridis I, Tavridou A, Christakidis D, Manolopoulos V. Presence of*CYP2C9**3allele increases risk for hypoglycemia in Type 2 diabetic patients treated with sulfonylureas. Pharmacogenomics. 2009;10(11):1781-1787.

130. Keskitalo J, Pasanen M, Neuvonen P, Niemi M. Different effects of the*ABCG2*c.421C>A SNP on the pharmacokinetics of fluvastatin, pravastatin and simvastatin. Pharmacogenomics. 2009;10(10):1617-1624.

131. Ansari M, Sauty G, Labuda M, Gagné V, Rousseau J, Moghrabi A et al. Polymorphism in multidrug resistance-associated protein gene 3 is associated with outcomes in childhood acute lymphoblastic leukemia. The Pharmacogenomics Journal. 2011;12(5):386-394.

132. PHARMGKB. rs1128503 [Internet]. PharmGKB. 2018 [cited 27 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166157288/clinicalAnnotation/1449165158

133. PHARMGKB. rs762551 [Internet]. PharmGKB. 2020 [cited 27 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166154850/clinicalAnnotation/1447983940

134. PHARMGKB. Annotation of rs2242480 [Internet]. PharmGKB. 2014 [cited 27 September 2020]. Available from: https://www.pharmgkb.org/variantAnnotation/1444935523

135. PHARMGKB. Annotation of rs1128503 [Internet]. PharmGKB. 2013 [cited 27 September 2020]. Available from: https://www.pharmgkb.org/variantAnnotation/981501792

136. Aquilante C, Kosmiski L, Bourne D, Bushman L, Daily E, Hammond K et al. Impact of the*CYP2C8**3polymorphism on the drug-drug interaction between gemfibrozil and pioglitazone. British Journal of Clinical Pharmacology. 2012;75(1):217-226.

137. Lane S, Al-Zubiedi S, Hatch E, Matthews I, Jorgensen A, Deloukas P et al. The population pharmacokinetics of R- and S-warfarin: effect of genetic and clinical factors. British Journal of Clinical Pharmacology. 2011;73(1):66-76.

138. PHARMGKB. Annotation of rs2032582 [Internet]. PharmGKB. 2020 [cited 27 September 2020]. Available from: https://www.pharmgkb.org/variantAnnotation/1447949684

139. PHARMGKB. Annotation of rs1057910 [Internet]. PharmGKB. 2008 [cited 27 September 2020]. Available from: https://www.pharmgkb.org/variantAnnotation/982033064

140. Steimer W, Zöpf K, von Amelunxen S, Pfeiffer H, Bachofer J, Popp J et al. Allele-Specific Change of Concentration and Functional Gene Dose for the Prediction of Steady-State Serum Concentrations of Amitriptyline and Nortriptyline in *CYP2C19* and *CYP2D6* Extensive and Intermediate Metabolizers. Clinical Chemistry. 2004;50(9):1623-1633.

141. Kim K, Park P, Park J. Effect of *ABCB1* (MDR1) haplotypes derived from G2677T/C3435T on the pharmacokinetics of amlodipine in healthy subjects. British Journal of Clinical Pharmacology. 2007;63(1):53-58.

142. PHARMGKB. rs4149056 [Internet]. PharmGKB. 2020 [cited 27 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166154579/clinicalAnnotation/655384011

143. PHARMGKB. rs12248560 [Internet]. PharmGKB. 2016 [cited 27 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166154147/clinicalAnnotation/655386913

144. PHARMGKB. rs4244285 [Internet]. PharmGKB. 2020 [cited 27 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166154053/clinicalAnnotation/1043858794

145. PHARMGKB. rs28371725 [Internet]. PharmGKB. 2015 [cited 27 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166156155/clinicalAnnotation/1183620242

146. PHARMGKB. rs1065852 [Internet]. PharmGKB. 2020 [cited 27 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166156062/clinicalAnnotation/1183624646

147. Chen L, Hong C, Chen E, Yee S, Xu L, Almof E et al. Genetic and epigenetic regulation of the organic cation transporter 3, *SLC22A3*. The Pharmacogenomics Journal. 2012;13(2):110-120.

148. Beermann B, Grind M, Lindstrom B. Pharmacokinetics of bendroflumethiazide. Clinical Pharmacology & Therapeutics. 1977;22(4):385-388.

149. Li M, Seiser E, Baldwin R, Ramirez J, Ratain M, Innocenti F et al. ABC transporter polymorphisms are associated with irinotecan pharmacokinetics and neutropenia. The Pharmacogenomics Journal. 2016;18(1):35-42.

150. DrugBank. Rosuvastatin [Internet]. Go.drugbank.com. 2020 [cited 27 September 2020]. Available from: https://go.drugbank.com/drugs/DB01098

151. PHARMGKB. Annotation of rs2231135 [Internet]. PharmGKB. 2015 [cited 27 September 2020]. Available from: https://www.pharmgkb.org/variantAnnotation/1447675109

152. PHARMGKB. rosuvastatin [Internet]. PharmGKB. 2017 [cited 27 September 2020]. Available from: https://www.pharmgkb.org/chemical/PA134308647/clinicalAnnotation/1154221922

153. Elliot D, Lewis B, Gillam E, Birkett D, Gross A, Miners J. Identification of the human cytochromes P450 catalysing the rate-limiting pathways of gliclazide elimination. British Journal of Clinical Pharmacology. 2007;64(4):450-457.

154. Inoue K, Yuasa H. Molecular Basis for Pharmacokinetics and Pharmacodynamics of Methotrexate in Rheumatoid Arthritis Therapy. Drug Metabolism and Pharmacokinetics. 2014;29(1):12-19.

155. Zhang Y, Li M, Tang J, Chen X. Pharmacokinetic and Pharmacodynamic Responses to Clopidogrel: Evidences and Perspectives. International Journal of Environmental Research and Public Health. 2017;14(3):301.

156. Dean L. Amitriptyline Therapy and *CYP2D6* and *CYP2C19* Genotype [Internet]. National Center for Biotechnology Information; 2017 [cited 27 September 2020]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK425165/

157. DrugBank. Valsartan [Internet]. Go.drugbank.com. 2020 [cited 28 September 2020]. Available from: https://go.drugbank.com/drugs/DB00177

158. Oswald S, Haenisch S, Fricke C, Sudhop T, Remmler C, Giessmannt et al. Intestinal expression of P-glycoprotein (*ABCB1*), multidrug resistance associated protein 2 (*ABCC2*), and uridine diphosphate–glucuronosyltransferase 1A1 predicts the disposition and modulates the effects of the cholesterol absorption inhibitor ezetimibe in humans. Clinical Pharmacology & Therapeutics. 2006;79(3):206-217.

159. Elens L, Vandercam B, Yombi J, Lison D, Wallemacq P, Haufroid V. Influence of host genetic factors on efavirenz plasma and intracellular pharmacokinetics in HIV-1-infected patients. Pharmacogenomics. 2010;11(9):1223-1234.

160. Thomsen R, Rasmussen H, Linnet K. In Vitro Drug Metabolism by Human Carboxylesterase 1: Focus on Angiotensin-Converting Enzyme Inhibitors. Drug Metabolism and Disposition. 2013;42(1):126-133.

161. Hamzic S, Kummer D, Milesi S, Mueller D, Joerger M, Aebi S et al. Novel Genetic Variants in Carboxylesterase 1 Predict Severe Early-Onset Capecitabine-Related Toxicity. Clinical Pharmacology & Therapeutics. 2017;102(5):796-804.

162. Pauli-Magnus C, Rekersbrink S, Klotz U, Fromm M. Interaction of omeprazole, lansoprazole and pantoprazole with P-glycoprotein. Naunyn-Schmiedeberg's Archives of Pharmacology. 2001;364(6):551-557.

163. PHARMGKB. rs9282564 [Internet]. PharmGKB. 2020 [cited 28 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166157435/variantAnnotation

164. Mimura Y, Yasujima T, Ohta K, Inoue K, Yuasa H. Functional identification of organic cation transporter 1 as an atenolol transporter sensitive to flavonoids. Biochemistry and Biophysics Reports. 2015;2:166-171.

165. Tarasova L, Kalnina I, Geldnere K, Bumbure A, Ritenberga R, Nikitina-Zake L et al. Association of genetic variation in the organic cation transporters OCT1, OCT2 and multidrug and toxin extrusion 1 transporter protein genes with the gastrointestinal side effects

and lower BMI in metformin-treated type 2 diabetes patients. Pharmacogenetics and Genomics. 2012;22(9):659-666.

166. DrugBank. Digoxin [Internet]. Go.drugbank.com. 2020 [cited 28 September 2020]. Available from: https://go.drugbank.com/drugs/DB00390

167. DrugBank. Nifedipine [Internet]. Go.drugbank.com. 2020 [cited 28 September 2020]. Available from: https://go.drugbank.com/drugs/DB01115

168. Ceppi F, Langlois-Pelletier C, Gagné V, Rousseau J, Ciolino C, Lorenzo S et al. Polymorphisms of the vincristine pathway and response to treatment in children with childhood acute lymphoblastic leukemia. Pharmacogenomics. 2014;15(8):1105-1116.

169. Meyer M, Seitz T, Brockmöller J, Tzvetkov M. Effects of genetic polymorphisms on the OCT1 and OCT2-mediated uptake of ranitidine. PLOS ONE. 2017;12(12):e0189521.

170. PHARMGKB. rs316019 [Internet]. PharmGKB. [cited 28 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166156913/variantAnnotation

171. Kiang T, Ho P, Anari M, Tong V, Abbott F, Chang T. Contribution of *CYP2C9*, *CYP2A6*, and *CYP2B6* to Valproic Acid Metabolism in Hepatic Microsomes from Individuals with the *CYP2C9**1/*1 Genotype. Toxicological Sciences. 2006;94(2):261-271.

172. PHARMGKB. Very Important Pharmacogene: *CYP2B6* [Internet]. PharmGKB. 2017 [cited 28 September 2020]. Available from: https://www.pharmgkb.org/vip/PA166169423

173. Sangkuhl K, Klein T, Altman R. Clopidogrel pathway. Pharmacogenetics and Genomics. 2010;1.

174. PHARMGKB. Annotation of rs3815583 [Internet]. PharmGKB. 2018 [cited 28 September 2020]. Available from: https://www.pharmgkb.org/variantAnnotation/827923850

175. Duman I. Role of Pharmacogenetics on Response to Statins: A Genotype-based Approach to Statin Therapy Outcome. Journal of Cardiology and Therapy. 2014;1(6).

176. LLerena A, Dorado P, Ramírez R, Calzadilla L, Peñas-LLedó E, Álvarez M et al. *CYP2D6* -1584C>G promoter polymorphism and debrisoquine ultrarapid hydroxylation in healthy volunteers. Pharmacogenomics. 2013;14(16):1973-1977.

177. Knauer M, Urquhart B, Meyer zu Schwabedissen H, Schwarz U, Lemke C, Leake B et al. Human Skeletal Muscle Drug Transporters Determine Local Exposure and Toxicity of Statins. Circulation Research. 2010;106(2):297-306.

178. Cha P, Mushiroda T, Zembutsu H, Harada H, Shinoda N, Kawamoto S et al. Single nucleotide polymorphism in *ABCG2* is associated with irinotecan-induced severe myelosuppression. Journal of Human Genetics. 2009;54(10):572-580.

179. DrugBank. Esomeprazole [Internet]. Go.drugbank.com. 2020 [cited 28 September 2020]. Available from: https://go.drugbank.com/drugs/DB00736

180. Hunfeld N, Touw D, Mathot R, Schaik R, Kuipers E. A comparison of the acid-inhibitory effects of esomeprazole and rabeprazole in relation to pharmacokinetics and *CYP2C19* polymorphism. Alimentary Pharmacology & Therapeutics. 2012;35(7):810-818.

181. Lou H, Chang C, Sheu M, Chen Y, Ho H. Optimal dose regimens of esomeprazole for gastric acid suppression with minimal influence of the *CYP2C19* polymorphism. European Journal of Clinical Pharmacology. 2008;65(1):55-64.

182. Scott S, Sangkuhl K, Shuldiner A, Hulot J, Thorn C, Altman R et al. PharmGKB summary: very important pharmacogene information for cytochrome P450, family 2, subfamily C, polypeptide 19. Pharmacogenetics and Genomics. 2012;22(2):159-165.

183. PHARMGKB. rs1128503 [Internet]. PharmGKB. [cited 28 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166157288/clinicalAnnotation

184. PHARMGKB. rs1045642 [Internet]. PharmGKB. [cited 28 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166157284/clinicalAnnotation

185. Venkatasubramanian R, Fukuda T, Niu J, Mizuno T, Chidambaran V, Vinks A et al. *ABCC3*andOCT1genotypes influence pharmacokinetics of morphine in children. Pharmacogenomics. 2014;15(10):1297-1309.

186. PHARMGKB. rs2231142 [Internet]. PharmGKB. 2019 [cited 28 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166156544/clinicalAnnotation/1447982582

187. Takada T, Ichida K, Matsuo H, Nakayama A, Murakami K, Yamanashi Y et al. ABCG2 Dysfunction Increases Serum Uric Acid by Decreased Intestinal Urate Excretion. Nucleosides, Nucleotides and Nucleic Acids. 2014;33(4-6):275-281.

188. Jabeen S, Holmboe L, Alnæs G, Andersen A, Hall K, Kristensen V. Impact of genetic variants of RFC1, DHFR and MTHFR in osteosarcoma patients treated with high-dose methotrexate. The Pharmacogenomics Journal. 2015;15(5):385-390.

189. Ferslew B, Köck K, Bridges A, Brouwer K. Role of Multidrug Resistance–Associated Protein 4 in the Basolateral Efflux of Hepatically Derived Enalaprilat. Drug Metabolism and Disposition. 2014;42(9):1567-1574.

190. den Hoed M, Lopez-Lopez E, te Winkel M, Tissing W, de Rooij J, Gutierrez-Camino A et al. Genetic and metabolic determinants of methotrexate-induced mucositis in pediatric acute lymphoblastic leukemia. The Pharmacogenomics Journal. 2014;15(3):248-254.

191. Jorgensen A, Al-Zubiedi S, Zhang J, Keniry A, Hanson A, Hughes D et al. Genetic and environmental factors determining clinical outcomes and cost of warfarin therapy: a prospective study. Pharmacogenetics and Genomics. 2009;19(10):800-812.

Chapter IV

192. Harper R, Ennis C, Sheridan B, Atkinson A, Johnston G, Bell P. Effects of low dose versus conventional dose thiazide diuretic on insulin action in essential hypertension. BMJ. 1994;309(6949):226-230.

193. PHARMGKB. rs2199939 [Internet]. PharmGKB. 2018 [cited 28 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166157575/clinicalAnnotation/1447952004

194. PHARMGKB. Atorvastatin/Lovastatin/Simvastatin Pathway, Pharmacokinetics [Internet]. PharmGKB. 2011 [cited 28 September 2020]. Available from: https://www.pharmgkb.org/pathway/PA145011109

195. Kitzmiller J, Mikulik E, Dauki A, Mukherjee C, Luzum J. Pharmacogenomics of statins: understanding susceptibility to adverse effects. Pharmacogenomics and Personalized Medicine. 2016;Volume 9:97-106.

196. Bakos É, Evers R, Sinkó E, Váradi A, Borst P, Sarkadi B. Interactions of the Human Multidrug Resistance Proteins MRP1 and MRP2 with Organic Anions. Molecular Pharmacology. 2000;57(4):760-768.

197. Becker M, Elens L, Visser L, Hofman A, Uitterlinden A, van Schaik R et al. Genetic variation in the *ABCC2* gene is associated with dose-decreases or switches to other cholesterol-lowering drugs during simvastatin and atorvastatin therapy. The Pharmacogenomics Journal. 2011;13(3):251-256.

198. PHARMGKB. rs2032582 [Internet]. PharmGKB. 2018 [cited 28 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166157316/clinicalAnnotation/1150414901

199. Uhr M, Tontsch A, Namendorf C, Ripke S, Lucae S, Ising M et al. Polymorphisms in the Drug Transporter Gene *ABCB1* Predict Antidepressant Treatment Response in Depression. Neuron. 2008;57(2):203-209.

200. Knauer M, Urquhart B, Meyer zu Schwabedissen H, Schwarz U, Lemke C, Leake B et al. Human Skeletal Muscle Drug Transporters Determine Local Exposure and Toxicity of Statins. Circulation Research. 2010;106(2):297-306.

201. Chen S, Villeneuve L, Jonker D, Couture F, Laverdière I, Cecchin E et al. *ABCC5* and ABCG1 polymorphisms predict irinotecan-induced severe toxicity in metastatic colorectal cancer patients. Pharmacogenetics and Genomics. 2015;25(12):573-583.

202. Keskitalo J, Zolk O, Fromm M, Kurkinen K, Neuvonen P, Niemi M. *ABCG2* Polymorphism Markedly Affects the Pharmacokinetics of Atorvastatin and Rosuvastatin. Clinical Pharmacology & Therapeutics. 2009;86(2):197-203.

203. Warren R, Smith R, Campalani E, Eyre S, Smith C, Barker J et al. Genetic Variation in Efflux Transporters Influences Outcome to Methotrexate Therapy in Patients with Psoriasis. Journal of Investigative Dermatology. 2008;128(8):1925-1929.

204. Jacobsen W, Kuhn B, Soldner A, Kirchner G, Sewing K, Kollman, P et al. Lactonization is the critical first step in the disposition of the 3-hydroxy-3-methylglutaryl-CoA reductase inhibitor atorvastatin. Drug Metab Dispos . 2000;28(11):1369-78.

205. Zhu Y, Wang F, Li Q, Zhu M, Du A, Tang W et al. Amlodipine Metabolism in Human Liver Microsomes and Roles of *CYP3A4*/5 in the Dihydropyridine Dehydrogenation. Drug Metabolism and Disposition. 2013;42(2):245-249.

206. FDA. CADUET (amlodipine besylate/atorvastatin calcium) Tablets [Internet]. 2009 [cited 28 September 2020]. Available from: https://www.accessdata.fda.gov/drugsatfda_docs/label/2009/021540s018lbl.pdf

207. Wang D, Guo Y, Wrighton S, Cooke G, Sadee W. Intronic polymorphism in *CYP3A4* affects hepatic expression and response to statin drugs. The Pharmacogenomics Journal. 2010;11(4):274-286.

208. Becker M, Visser L, van Schaik R, Hofman A, Uitterlinden A, Ch. Stricker B. Influence of genetic variation in *CYP3A4* and *ABCB1* on dose-decrease or switching during simvastatin and atorvastatin therapy. Pharmacoepidemiology and Drug Safety. 2009;19(1):75-81.

209. Chen C, Wang S, Tsou H, Ho I, Tian J, Yu C et al. Genetic polymorphisms in CYP3A4 are associated with withdrawal symptoms and adverse reactions in methadone maintenance patients. Pharmacogenomics. 2011;12(10):1397-1406.

210. DrugBank. Simvastatin [Internet]. Go.drugbank.com. 2020 [cited 28 September 2020]. Available from: https://go.drugbank.com/drugs/DB00641

211. Kitzmiller J, Mikulik E, Dauki A, Mukherjee C, Luzum J. Pharmacogenomics of statins: understanding susceptibility to adverse effects. Pharmacogenomics and Personalized Medicine. 2016;Volume 9:97-106.

212. DrugBank. Warfarin [Internet]. Go.drugbank.com. 2020 [cited 28 September 2020]. Available from: https://go.drugbank.com/drugs/DB00682

213. Omar M, Wilson J. FDA Adverse Event Reports on Statin-Associated Rhabdomyolysis. Annals of Pharmacotherapy. 2002;36(2):288-295.

214. PHARMGKB. rs4148386 [Internet]. PharmGKB. 2018 [cited 28 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166154047/clinicalAnnotation/1448098947

215. Andersson M, Mannheimer B, Lindh J. The effect of simvastatin on warfarin anticoagulation: a Swedish register-based nationwide cohort study. European Journal of Clinical Pharmacology. 2019;75(10):1387-1392.

216. Andersson M, Eliasson E, Lindh J. A clinically significant interaction between warfarin and simvastatin is unique to carriers of the*CYP2C9**3allele. Pharmacogenomics. 2012;13(7):757-762.

217. PHARMGKB. rs3814637 [Internet]. PharmGKB. 2016 [cited 28 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166154042/clinicalAnnotation/1448100296

218. Horikiri Y, Suzuki T, Mizobe M. Pharmacokinetics and Metabolism of Bisoprolol Enantiomers in Humans. Journal of Pharmaceutical Sciences. 1998;87(3):289-294.

219. Hukkanen J, Puurunen J, Hyötyläinen T, Savolainen M, Ruokonen A, Morin-Papunen L et al. The effect of atorvastatin treatment on serum oxysterol concentrations and cytochrome P450 3A4 activity. British Journal of Clinical Pharmacology. 2015;80(3):473-479.

220. Kobie J, Guo Z, Cho C, Menzel K, McCrea J, Blanchard R et al. Pharmacogenetic Analysis of OATP1B1, UGT1A1 , and BCRP Variants in Relation to the Pharmacokinetics of Letermovir in Previously Conducted Clinical Studies. The Journal of Clinical Pharmacology. 2019.

221. Chigutsa E, Visser M, Swart E, Denti P, Pushpakom S, Egan D et al. The*SLCO1B1*rs4149032 Polymorphism Is Highly Prevalent in South Africans and Is Associated with Reduced Rifampin Concentrations: Dosing Implications. Antimicrobial Agents and Chemotherapy. 2011;55(9):4122-4127.

222. Low S, Kiyotani K, Mushiroda T, Daigo Y, Nakamura Y, Zembutsu H. Association study of genetic polymorphism in *ABCC4* with cyclophosphamide-induced adverse drug reactions in breast cancer patients. Journal of Human Genetics. 2009;54(10):564-571.

223. DrugBank. Omeprazole [Internet]. Go.drugbank.com. 2020 [cited 28 September 2020]. Available from: https://go.drugbank.com/drugs/DB00338

224. DrugBank. Lansoprazole [Internet]. Go.drugbank.com. 2020 [cited 28 September 2020]. Available from: https://go.drugbank.com/drugs/DB00448

225. DrugBank. Atorvastatin [Internet]. Go.drugbank.com. 2020 [cited 28 September 2020]. Available from: https://go.drugbank.com/drugs/DB01076

226. Katoh M, Nakajima M, Shimada N, Yamazaki H, Yokoi T. Inhibition of human cytochrome P450 enzymes by 1,4-dihydropyridine calcium antagonists: prediction of in vivo drug-drug interactions. European Journal of Clinical Pharmacology. 2000;55(11-12):843-852.

227. PHARMGKB. Omeprazole [Internet]. PharmGKB. 2020 [cited 28 September 2020]. Available from: https://www.pharmgkb.org/chemical/PA450704/clinicalAnnotation/1183629924

228. PHARMGKB. Lansoprazole [Internet]. PharmGKB. 2020 [cited 28 September 2020]. Available from: https://www.pharmgkb.org/chemical/PA450180/clinicalAnnotation/1183629580

229. Brophy J, Babapulle M, Costa V, Rinfret S. A pharmacoepidemiology study of the interaction between atorvastatin and clopidogrel after percutaneous coronary intervention. American Heart Journal. 2006;152(2):263-269.

230. Chew S, Sandanaraj E, Singh O, Chen X, Tan E, Lim W et al. Influence of *SLCO1B3* haplotype-tag SNPs on docetaxel disposition in Chinese nasopharyngeal cancer patients. British Journal of Clinical Pharmacology. 2012;73(4):606-618.

231. Barkas F, Elisaf M, Rizos C, Klouras E, Kostapanos M, Liberopoulos E. Proton pump inhibitors and statins: a possible interaction that favors low-density lipoprotein cholesterol reduction?. Hippokratia. 2015;19(4):332–337.

232. FIiegenabsum M, Dasilveira F, Vandersand C, Vandersand L, Ferreira M, Pires R et al. The role of common variants of *ABCB1*, *CYP3A4*, and *CYP3A5* genes in lipid-lowering efficacy and safety of simvastatin treatment. Clinical Pharmacology & Therapeutics. 2005;78(5):551-558.

233. Becker M, Visser L, van Schaik R, Hofman A, Uitterlinden A, Stricker B. Common genetic variation in the*ABCB1*gene is associated with the cholesterol-lowering effect of simvastatin in males. Pharmacogenomics. 2009;10(11):1743-1751.

234. Dahlin A, Wittwer M, de la Cruz M, Woo J, Bam R, Scharen-Guivel V et al. A pharmacogenetic candidate gene study of tenofovir-associated Fanconi syndrome. Pharmacogenetics and Genomics. 2015;25(2):82-92.

235. Izzedine H, Hulot J, Villard E, Goyenvalle C, Dominguez S, Ghosn J et al. Association between*ABCC2*Gene Haplotypes and Tenofovir-Induced Proximal Tubulopathy. The Journal of Infectious Diseases. 2006;194(11):1481-1491.

236. PHARMGKB. rs4149057 [Internet]. PharmGKB. 2018 [cited 28 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166154580/clinicalAnnotation/1043880701

237. Nishio S, Watanabe H, Kosuge K, Uchida S, Hayashi H, Ohashi K. Interaction between Amlodipine and Simvastatin in Patients with Hypercholesterolemia and Hypertension. Hypertension Research. 2005;28(3):223-227.

238. MHRA. MHRA recommendations on simvastatin interactions: What are the implications for patients taking amlodipine? [Internet]. 2012 [cited 28 September 2020]. Available from: https://mm.wirral.nhs.uk/document_uploads/alerts/NWMICsummarysimvaamlodipineinteractionSep12.pdf

239. Kim K, Park P, Park J. Effect of *ABCB1* (MDR1) haplotypes derived from G2677T/C3435T on the pharmacokinetics of amlodipine in healthy subjects. British Journal of Clinical Pharmacology. 2007;63(1):53-58.

240. Darvari R, Boroujerdi M. Concentration dependency of modulatory effect of amlodipine on P-glycoprotein efflux activity of doxorubicin - a comparison with tamoxifen. Journal of Pharmacy and Pharmacology. 2004;56(8):985-991.

241. PHARMGKB. rs3842 [Internet]. PharmGKB. 2020 [cited 28 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166157239/clinicalAnnotation/1184133889

242. PHARMGKB. rs1045642 [Internet]. PharmGKB. 2020 [cited 28 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166157284/clinicalAnnotation/1142234402

243. Rodrigues A, Perin P, Purim S, Silbiger V, Genvigir F, Willrich M et al. Pharmacogenetics of OATP Transporters Reveals That *SLCO1B1* c.388A>G Variant Is Determinant of Increased Atorvastatin Response. International Journal of Molecular Sciences. 2011;12(9):5815-5827.

244. Fu Q, Li Y, Gao Y, Yang S, Lu P, Jia M et al. Lack of association between *SLCO1B1* polymorphism and the lipid-lowering effects of atorvastatin and simvastatin in Chinese individuals. European Journal of Clinical Pharmacology. 2012;69(6):1269-1274.

Chapter V

245. Ioannidis J, Tarone R, McLaughlin J. The False-positive to False-negative Ratio in Epidemiologic Studies. Epidemiology. 2011;22(4):450-456.

246. DrugBank. Quinine [Internet]. Go.drugbank.com. 2020 [cited 28 September 2020]. Available from: https://go.drugbank.com/drugs/DB00468

247. Lee S, Jang Y, Cha E, Kim H, Lee S, Shin J. A haplotype of *CYP2C9* associated with warfarin sensitivity in mechanical heart valve replacement patients. British Journal of Clinical Pharmacology. 2010;70(2):213-221.

248. Sortica V, Lindenau J, Cunha M, Ohnishi M, Ventura A, Ribeiro-dos-Santos A et al. The effect of SNPs in CYP450 in chloroquine/primaquine Plasmodium vivax malaria treatment. Pharmacogenomics. 2016;17(17):1903‑1911.

249. Polonikov A, Kharchenko A, Bykanova M, Sirotina S, Ponomarenko I, Bocharova A et al. Polymorphisms of CYP2C8 , CYP2C9 and CYP2C19 and risk of coronary heart disease in Russian population. Gene. 2017;627:451-459

250. FDA. CARDURA® (doxazosin mesylate) tablets for oral use [Internet]. 2016 [cited 28 September 2020]. Available from: https://www.accessdata.fda.gov/drugsatfda_docs/label/2016/019668s028lbl.pdf

251. Ansari M, Sauty G, Labuda M, Gagné V, Rousseau J, Moghrabi A et al. Polymorphism in multidrug resistance-associated protein gene 3 is associated with outcomes in childhood acute lymphoblastic leukemia. The Pharmacogenomics Journal. 2011;12(5):386-394.

252. DrugBank. Ramipril [Internet]. Go.drugbank.com. 2020 [cited 28 September 2020]. Available from: https://go.drugbank.com/drugs/DB00178

253. PharmVar. *CYP2D6* [Internet]. Pharmvar.org. 2020 [cited 28 September 2020]. Available from: https://www.pharmvar.org/gene/*CYP2D6*

254. Kamal Z, Nordstad M, Molden E. A 60-year-old man with heart failure, dry cough and elevated INR values [Internet]. Tidsskrift for Den norske legeforening. 2008 [cited 6 February 2021]. Available from: https://tidsskriftet.no/en/2008/08/case-reports/60-year-old-man-heart-failure-dry-cough-and-elevated-inr-values

255. DrugBank. Amlodipine [Internet]. Go.drugbank.com. 2020 [cited 28 September 2020]. Available from: https://go.drugbank.com/drugs/DB00381#label-reference

256. Ansari M, Sauty G, Labuda M, Gagné V, Laverdière C, Moghrabi A et al. Polymorphisms in multidrug resistance-associated protein gene 4 is associated with outcome in childhood acute lymphoblastic leukemia. Blood. 2009;114(7):1383-1386.

257. Sangkuhl K, Shuldiner A, Klein T, Altman R. Platelet aggregation pathway. Pharmacogenetics and Genomics. 2011;21(8):516-521.

258. Verschuren J, Boden H, Wessels J, van der Hoeven B, Trompet S, Heijmans B et al. Value of platelet pharmacogenetics in common clinical practice of patients with ST-segment elevation myocardial infarction. International Journal of Cardiology. 2013;167(6):2882-2888.

259. Li X, Ma N, Li X, Wang B, Sun S, Gao F et al. Association of PON1, P2Y12 and COX1 with Recurrent Ischemic Events in Patients with Extracranial or Intracranial Stenting. PLOS ONE. 2016;11(2):e0148891.

260. FDA. Procardia XL (nifedipine) Extended Release Tablets For Oral Use [Internet]. 2014 [cited 28 September 2020]. Available from: https://www.accessdata.fda.gov/drugsatfda_docs/label/2015/019684s029lbl.pdf

261. PHARMGKB. rs152023 [Internet]. PharmGKB. [cited 28 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166162356/variantAnnotation

262. DrugBank. Valproic acid [Internet]. Go.drugbank.com. 2020 [cited 28 September 2020]. Available from: https://go.drugbank.com/drugs/DB00313

263. Noai M, Soraoka H, Kajiwara A, Tanamachi Y, Oniki K, Nakagawa K et al. Cytochrome P450 2C19polymorphisms and valproic acid-induced weight gain. Acta Neurologica Scandinavica. 2015;133(3):216-223.

264. WANG S, Li J, Song M, Yan P, Ju X, Liu J et al. Effect of Genetic Polymorphisms of *CYP2C19* on the Steady-state Serum Concentrations of Valproic Acid in Chinese Han Patients With Schizophrenia. Annals of General Psychiatry. 2020;.

265. Gong L, Goswami S, Giacomini K, Altman R, Klein T. Metformin pathways: pharmacokinetics and pharmacodynamics. Pharmacogenetics and Genomics. 2012;22(11):820-827.

266. Dujic T, Causevic A, Bego T, Malenica M, Velija-Asimi Z, Pearson E et al. Organic cation transporter 1 variants and gastrointestinal side effects of metformin in patients with Type 2 diabetes. Diabetic Medicine. 2015;33(4):511-514.

267. Hemauer S, Patrikeeva S, Nanovskaya T, Hankins G, Ahmed M. Role of human placental apical membrane transporters in the efflux of glyburide, rosiglitazone, and metformin. American Journal of Obstetrics and Gynecology. 2010;202(4):383.e1-383.e7.

268. Hodges L, Markova S, Chinn L, Gow J, Kroetz D, Klein T et al. Very important pharmacogene summary: *ABCB1* (MDR1, P-glycoprotein). Pharmacogenetics and Genomics. 2011;21(3):152-161.

269. The electronic medicines compendium (emc). Nicorandil 10mg Tablets - Patient Information Leaflet (PIL) - (emc) [Internet]. Medicines.org.uk. [cited 28 September 2020]. Available from: https://www.medicines.org.uk/emc/product/652/pil

270. Cheng H, Harris R. Cyclooxygenases, the Kidney, and Hypertension. Hypertension. 2004;43(3):525-530.

271. Lopez-Lopez E, Ballesteros J, Piñan M, Sanchez de Toledo J, Garcia de Andoin N, Garcia-Miguel P et al. Polymorphisms in the methotrexate transport pathway: a new tool for MTX plasma level prediction in pediatric acute lymphoblastic leukemia. Pharmacogenetics and Genomics. 2013;23(2):53-61.

272. Inoue K, Yuasa H. Molecular Basis for Pharmacokinetics and Pharmacodynamics of Methotrexate in Rheumatoid Arthritis Therapy. Drug Metabolism and Pharmacokinetics. 2014;29(1):12-19.

273. The electronic medicines compendium (emc). Bisoprolol 2.5mg film coated tablet - Summary of Product Characteristics (SmPC) - (emc) [Internet]. Medicines.org.uk. [cited 28 September 2020]. Available from: https://www.medicines.org.uk/emc/product/6126/smpc

274. Wang D, Jonker J, Kato Y, Kusuhara H, Schinkel A, Sugiyama Y. Involvement of Organic Cation Transporter 1 in Hepatic and Intestinal Distribution of Metformin. Journal of Pharmacology and Experimental Therapeutics. 2002;302(2):510-515.

275. Becker M, Visser L, van Schaik R, Hofman A, Uitterlinden A, Stricker B. Genetic variation in the organic cation transporter 1 is associated with metformin response in patients with diabetes mellitus. The Pharmacogenomics Journal. 2009;9(4):242-247.

276. Zhang J, Jiao K, Liu J, Xia Y. Metformin reverses the resistance mechanism of lung adenocarcinoma cells that knocks down the Nrf2 gene. Oncology Letters. 2018;16(5):6071–6080.

Chapter VI

277. FDA. QUALAQUIN (quinine sulfate CAPSULES USP, 324 mg) [Internet]. 2008 [cited 29 September 2020]. Available from: https://www.accessdata.fda.gov/drugsatfda_docs/label/2008/021799s008lbl.pdf

278. Achan J, Talisuna A, Erhart A, Yeka A, Tibenderana J, Baliraine F et al. Quinine, an old anti-malarial drug in a modern world: role in the treatment of malaria. Malaria Journal. 2011;10(1).

279. Allen R, KirbyY K. Nocturnal Leg Cramps. Am Fam Physician. 2012;86(4):350-355.

280. FDA. New risk management plan and patient Medication Guide for Qualaquin [Internet]. U.S. Food and Drug Administration. 2017 [cited 29 September 2020]. Available from: https://www.fda.gov/drugs/postmarket-drug-safety-information-patients-and-providers/fda-drug-safety-communication-new-risk-management-plan-and-patient-medication-guide-qualaquin

281. Medicines and Healthcare products Regulatory Agency (MHRA). Quinine: not to be used routinely for nocturnal leg cramps [Internet]. GOV.UK. 2014 [cited 29 September 2020]. Available from: https://www.gov.uk/drug-safety-update/quinine-not-to-be-used-routinely-for-nocturnal-leg-cramps

282. Looareesuwan S, Charoenpan P, Ho M, White N, Karbwang J, Bunnag D et al. Fatal Plasmodium falciparum Malaria after an Inadequate Response to Quinine Treatment. Journal of Infectious Diseases. 1990;161(3):577-580.

283. Newton P, Ward S, Angus B, Chierakul W, Dondorp A, Ruangveerayuth R et al. Early treatment failure in severe malaria resulting from abnormally low plasma quinine concentrations. Trans R Soc Trop Med Hyg . 2005;100(2):184-6.

284. Adegbola A, Soyinka J, Adeagbo B, Igbinoba S, Nathaniel T. Alteration of the Disposition of Quinine in Healthy Volunteers After Concurrent Ciprofloxacin Administration. American Journal of Therapeutics. 2016;23(2):e398-e404.

285. PHARMGKB. rs4918758-Allele Frequencies [Internet]. PharmGKB. [cited 29 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166154067/overview

286. Kim J, Nguyen H. Doxazosin [Internet]. Ncbi.nlm.nih.gov. 2020 [cited 29 November 2020]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK557459/

287. NICE. Hypertension in adults: diagnosis and management [Internet]. Guidelines. 2020 [cited 29 November 2020]. Available from: https://www.guidelines.co.uk/cardiovascular/nice-hypertension-guideline/454934.article

288. PHARMGKB. rs9895420-Allele Frequencies [Internet]. PharmGKB. 2020 [cited 29 November 2020]. Available from: https://www.pharmgkb.org/variant/PA166155237

289. Fares H, DiNicolantonio J, O'Keefe J, Lavie C. Amlodipine in hypertension: a first-line agent with efficacy for improving blood pressure and patient outcomes. Open Heart. 2016;3(2).

290. FDA. Amlodipine Besylate Tablets [Internet]. 2007 [cited 29 September 2020]. Available from: https://www.accessdata.fda.gov/drugsatfda_docs/label/2007/019787s042lbl.pdf

291. Makani H, Bangalore S, Romero J, Htyte N, Berrios R, Makwana H et al. Peripheral edema associated with calcium channel blockers: incidence and withdrawal rate--a meta-analysis of randomized trials. J Hypertens. 2011;29(7):1270-80.

292. Bhatnagar V, Garcia E, O'Connor D, Brophy V, Alcaraz J, Richard E et al. *CYP3A4* and *CYP3A5* polymorphisms and blood pressure response to amlodipine among African-American men and women with early hypertensive renal disease. Am J Nephrol . 2010;31(2):95-103.

293. Sychev D, Shikh N, Morozova T, Grishina E, Ryzhikova K, Malova E. Effects of *ABCB1* rs1045642 polymorphisms on the efficacy and safety of amlodipine therapy in Caucasian patients with stage I–II hypertension. Pharmgenomics Pers Med. 2018;11:157–165.

294. PHARMGKB. rs868853-Allele Frequencies [Internet]. PharmGKB. [cited 29 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166154687

295. World Health Organization (WHO). Cardiovascular diseases [Internet]. Who.int. [cited 29 September 2020]. Available from: https://www.who.int/health-topics/cardiovascular-diseases/#tab=tab_2

296. Sangkuhl K, Shuldiner A, Klein T, Altman R. Platelet aggregation pathway. Pharmacogenet Genomics. 2011;21(8):516–521.

297. Chien L, Liu H, Chiou H, Chi N. Efficacy and safety of clopidogrel and aspirin do not differ in patients with stable ischemic stroke. J Chin Med Assoc. 2020;83(7):651-656.

298. CAPRIE Steering Committee. A randomised, blinded, trial of clopidogrel versus aspirin in patients at risk of ischaemic events (CAPRIE). CAPRIE Steering Committee. Lancet. 1996;348(9038):1329-39.

299. Ray S. Clopidogrel resistance: The way forward. Indian Heart J. 2014;66(5):530–534.

300. Floyd C, Ferro A. Antiplatelet drug resistance: Molecular insights and clinical implications. Prostaglandins & Other Lipid Mediators. 2015;120:21-27.

301. Michelson A, Frelinger A, Furman M. Resistance to antiplatelet drugs. European Heart Journal Supplements. 2006;8(suppl_G):G53-G58.

302. Patti G, Micieli G, Cimminiello C, Bolognese L. The Role of Clopidogrel in 2020: A Reappraisal. Cardiovasc Ther. 2020.

303. Jneid H, Anderson J, Wright S, Adams C, Bridges C, Casey D et al. 2012 ACCF/AHA Focused Update of the Guideline for the Management of Patients With Unstable Angina/Non–ST-Elevation Myocardial Infarction (Updating the 2007 Guideline and Replacing the 2011 Focused Update): A Report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. The American College of Cardiology. 2012;60(7):645-681.

304. Germaine C, Bogaty P, Boyer L, Hanley J, Engert J, Brophy J. Genetic Polymorphisms and the Cardiovascular Risk of Non-Steroidal Anti-Inflammatory Drugs. The American Journal of Cardiology. 2010;105(12):1740-1745.

305. PHARMGKB. rs12353214-Allele Frequencies [Internet]. PharmGKB. 2020 [cited 29 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166157837

306. Jiang X, Samant, S, Lesko L, Schmidt* S. Clinical Pharmacokinetics and Pharmacodynamics of Clopidogrel. Clin Pharmacokinet. 2015;54(2):147–166.
307. Tucker W, Sankar P, Kariyanna P. StatPearls-Selective Beta-1-Blockers [Internet]. NCBI; 2020 [cited 29 September 2020]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK499982/
308. Wong G, Boyda H, Wright J. Blood pressure lowering efficacy of beta-1 selective beta blockers for primary hypertension. Cochrane Database Syst Rev. 2016;2016(3).
309. Chavey W. The Importance of Beta Blockers in the Treatment of Heart Failure. Am Fam Physician. 2000;62(11):2453-2462.
310. Kezerashvili A, Marzo K, Leon J. Beta Blocker Use After Acute Myocardial Infarction in the Patient with Normal Systolic Function: When is it "Ok" to Discontinue?. Curr Cardiol Rev. 2012;8(1):77–84.
311. Wu T, Chen X, Deng L. Beta-blockers for unstable angina. Cochrane Database Syst Rev. 2017;2017(11).
312. Wee W, Burns K, Bett N. Medical management of chronic stable angina. Aust Prescr. 2015;38(4):131–136.
313. Bazroon A, Alrashidi N. StatPearls-Bisoprolol [Internet]. NCBI; 2020 [cited 29 September 2020]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK551623/
314. FDA. Zebeta (Bisiprolol fumarate) TABLETS - bisoprolol fumarate tablet [Internet]. 2007 [cited 29 September 2020]. Available from: https://www.accessdata.fda.gov/drugsatfda_docs/label/2007/019982s014lbl.pdf
315. PHARMGKB. rs9516519-Allele Frequencies [Internet]. PharmGKB. 2020 [cited 29 September 2020]. Available from: https://www.pharmgkb.org/variant/PA166154734
316. Mega J, Close S, Wiviott S, Shen L, Hockett R, Brandt J et al. Cytochrome p-450 polymorphisms and response to clopidogrel. N Engl J Med. 2009;360(4):354-62.
317. Slamon D, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. N Engl J Med. 2001;344(1):783-92.
318. Saudi Genome Project Team. The Saudi Human Genome Program: An oasis in the desert of Arab medicine is providing clues to genetic disease. IEEE Pulse. 2015;6(6):22-6.
319. Saudi Genome. Saudi Human Genome Program [Internet]. Shgp.sa. 2020 [cited 29 November 2020]. Available from: https://shgp.sa/#home

This page was left intentionally

Appendices:

Supplementary material 1:

The Pharmacogenomics Journal (2020) 20:355–366
https://doi.org/10.1038/s41397-019-0122-0

REVIEW ARTICLE

Readers' Choice: The best of *The Pharmacogenomics Journal* 2019

Drug–drug–gene interactions and adverse drug reactions

Mustafa Adnan Malki[1] · Ewan Robert Pearson[1]

Received: 5 November 2018 / Revised: 11 November 2019 / Accepted: 17 November 2019 / Published online: 3 December 2019

Abstract
The economic and health burden caused by adverse drug reactions has increased dramatically in the last few years. This is likely to be mediated by increasing polypharmacy, which increases the likelihood for drug–drug interactions. Tools utilized by healthcare practitioners to flag potential adverse drug reactions secondary to drug–drug interactions ignore individual genetic variation, which has the potential to markedly alter the severity of these interactions. To date there have been limited published studies on impact of genetic variation on drug–drug interactions. In this review, we establish a detailed classification for pharmacokinetic drug–drug–gene interactions, and give examples from the literature that support this approach. The increasing availability of real-world drug outcome data linked to genetic bioresources is likely to enable the discovery of previously unrecognized, clinically important drug–drug–gene interactions.

Introduction

It was previously and alarmingly reported that adverse drug reactions (ADRs) represent the fourth leading cause of death in the USA [1]. A recent review (2015) showed that 3.6% of patients were admitted to hospitals in Europe due to ADRs and 10% of patients developed side effects during their in-patient stay [2]. The latest report issued by MiDatabank in cooperation with the Medicines and Healthcare Products Regulatory Agency, shows an increasing trend in the number of reported ADRs in the period between 2011 and 2016 across the UK [3]. It has also been estimated that ADRs alone cost the NHS £770 M annually [4]. Nonsteroidal antiinflammatory drugs, diuretics, anticoagulants, and antiplatelets have been recognized to be the major culprits, with prescribing errors being major contributors to medication-related adverse events [5]. The chance of these errors increases when patients undergo multiple treatments; a situation that is highly prevalent in elderly patients [6]. There are a number of factors that influence the occurrence of ADRs secondary to drug–drug interactions, such as age, renal function, and other comorbidities. In addition, genetic variation is likely to play a crucial role in the development of ADRs. For example, when only considering genetic polymorphisms in three drug metabolizing enzymes (cytochrome P450 2C9 (CYP2C9), CYP2C19, and CYP2D6), 15% of the ADRs were due to drug–gene interactions, and 19% were due to drug–drug–gene interactions [7]. Incorporation of these gene variants increased the number of predicted clinically critical drug interactions by ~51% [7]. Given the large number of genes involved in drug metabolism and transport, we cannot underestimate the importance of genetic variation in contributing to potential for clinically critical ADRs.

Following the recent advances in pharmacogenomics, the traditional view of drug–drug interactions needs to be modified to include genetic variation. To date the literature on drug–drug–gene interactions (DDGIs) is limited, with only one previous review evaluating the impact of CYP2C9, C19, and 2D6 variants [8]. In this review, we attempt to provide an in-depth framework for the classification of pharmacokinetic DDGIs caused by different mechanisms, and their potential impact to increase clinically critical drug interactions in the context of the polypharmacy seen in modern medicine today.

Drug–drug–gene interactions

DDGIs can be divided into three main categories: inhibitory interactions, induction interactions, and phenoconversion

Supplementary information The online version of this article (https://doi.org/10.1038/s41397-019-0122-0) contains supplementary material, which is available to authorized users.

✉ Ewan Robert Pearson
e.z.pearson@dundee.ac.uk

[1] Population Health & Genomics, School of Medicine, University of Dundee, Dundee, UK

interactions. Inhibitory and induction interactions can be defined as any interactions that affect the victim drug's pharmacokinetics (PK) to increase or reduce concentrations of the drug, respectively. Induction or inhibition can occur either with the administration of a perpetrator drug that alters the victim drug metabolism or transport, or with the presence of loss- or gain-of-function (LOF or GOF) genetic variants that alter function of enzymes that alter metabolism or transport of the victim drug, or the combination of both. A DDGI can be thought of as a double hit—whereby the genetic variant and the perpetrator drug combine to act on transporter or metabolism pathways to greatly alter drug concentrations. It is also possible to see phenoconversion—where the interacting drug effect and the genotype have opposing effects, resulting in a temporary phenotype shift e.g. neutralizing/reversing the effect of a GOF genotype when an inhibitory drug is prescribed. In this review we describe, with examples, different cases of interactions under each of the above three categories, focusing initially on metabolizing enzymes, before considering drug transporters.

Drug–drug-metabolizing enzyme gene interactions (DDMEGIs)

Inhibitory interactions

Inhibitory effects of drugs and genotype can alter substrate metabolism by both drug and genotype impacting on the same metabolizing enzyme, or on two distinct routes of metabolism.

In general, poor metabolizers are expected to experience the highest substrate drug plasma concentration, compared with other genotypes, when co-treated with inhibitors. For example, co-administration of simvastatin (a CYP2C9 inhibitor) with warfarin (CYP2C9 substrate) has been shown to reduce warfarin dosage requirements in CYP2C9*3 carriers with a greater percentage as compared with noncarriers (29% vs 5% respectively) [9]. A similar conclusion has been reported with celecoxib (Supplementary Table 1) [10]. The inhibitory effect of drug and genotype is not always additive—genetically poor metabolizers may have only limited further enzyme inhibition by administration of an inhibitory drug. For instance, a statistically significant elevation in rabeprazole (a CYP2C19 substrate) plasma levels was observed in both normal metabolizers and heterozygous genotype carriers after treatment with fluvoxamine (a CYP2C19 inhibitor) while no additional clinically significant elevation was detected with poor metabolizers who have already experienced the highest rabeprazole plasma levels [11]. A similar scenario is seen with other examples (Supplementary Table 1) [12–15].

Where a drug is metabolized by two or more CYP enzymes, then inhibition of one of these enzymes alone (by drug or genotype) may have minimal effect, due to redundancy of the pathways. However, if a genotype and interacting drug affect these different routes of metabolism, then the interaction may be very large. For example, it has been observed that for voriconazole (a CYP2C19 and CYP3A4 substrate) bioavailability is increased markedly (~5.6-fold) in patients who have reduced CYP2C19 activity and are administered with atazanavir or ritonavir (potent CYP3A4 inhibitors) [16]. A similar scenario can be noted with other examples (Supplementary Table 1) [17–19].

Prodrugs, on the other hand, require the function of certain CYPs to be therapeutically active, and in these cases the effect is the opposite to that described above. Clopidogrel, for example, is activated by CYP1A2, CYP2B6, CYP3A4, CYP2C9, and CYP2C19 [20]. Carriers of LOF variants in one or more of these genes and co-administered with their inhibitors are at increased risk for treatment resistance. For instance, carriers of CYP2C19*2 and/or *3 alleles who are treated with clopidogrel and proton pump inhibitors (CYP2C19 inhibitors) were observed to be more likely to have reduced clopidogrel efficacy; the addition of a third risk factor (e.g., calcium channel blockers (CYP3A4 inhibitors)) was also correlated with a greater reduction in efficacy of clopidogrel [21, 22].

Figure 1 shows the predicted changes of plasma levels of active drugs and active metabolites of prodrugs with and without the presence of inhibitors and/or LOF variants.

Induction interactions

Increased metabolism of active drugs by an enzyme inducer or GOF variant will result in reduced efficacy of the victim drug. For example, when voriconazole (a CYP2C19 substrate) is co-prescribed with carbamazepine (CYP2C19 inducer) the voriconazole dose is usually increased to overcome this increased metabolism. In a case report, therapeutic concentrations of voriconazole were not achieved, as the patient carried two GOF CYP2C19 *17 variants [23].

The opposite effect is seen with prodrugs. Increased metabolism by an enzyme inducing drug or GOF variant, will result in high plasma levels of active metabolites leading to increased side effects and/or efficacy. Thus, patients carrying CYP2C19*17 GOF variants have increased conversion of clopidogrel to active metabolites resulting in reduced cardiovascular events and/or increased bleeding episodes [24–33]. Co-administration of an inducer of CYP1A2, CYP2C9, and/or CYP3A4 would be expected to result in greater efficacy of clopidogrel, with increased risk of bleeding, however no studies have been published to establish this.

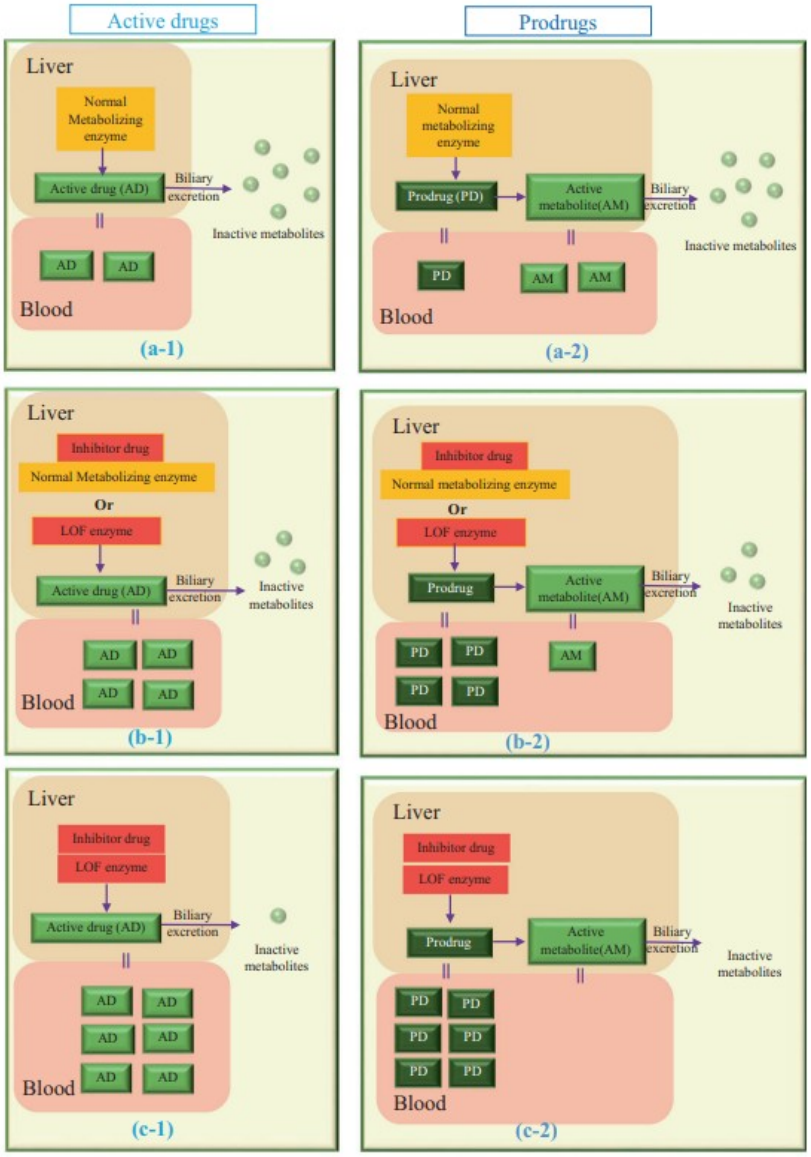


Fig. 1 The predicted active drug/active metabolites of prodrugs plasma levels and biliary excretion changes without or with the presence of inhibitors or LOF variants or both on metabolizing enzymes. The predicted active drug/active metabolites of prodrugs plasma levels and biliary excretion changes without (**a-1/a-2**) or with the presence of inhibitors or LOF variants (**b-1/b-2**) or both (**c-1/c-2**) on metabolizing enzymes. **a-1/a-2** represent the normal scenario with no interacting drug or genetic variant. In **b-1/b-2** either an inhibitory drug or loss-of-function variant (LOF) in the metabolizing enzyme, results in reduced metabolism to inactive metabolites, and increased (**b-1**)/decreased (**b-2**) active drug in the systemic circulation. In **c-1/c-2** the presence of inhibitory drug and the LOF genetic variant combine to produce greater increase (**c-1**)/decrease (**c-2**) in the systemic concentration of active drug

Figure 2 shows the predicted changes of plasma levels of active drugs and active metabolites of prodrugs with and without the presence of inducers and/or GOF variants.

Phenoconversion interactions

As described above, a temporary phenotype shift can be seen when the perpetrator drug and genetic effect are opposed. For example, the presence of reduced function CYP2C9 variants results in reduced tolbutamide (a CYP2C9 substrate) metabolism, yet co-treatment with rifampicin (a CYP2C9 inducer) in these patients reverses this genetic effect resulting in a twofold increase in tolbutamide clearance [34]. Conversely, proton pump inhibitors (CYP2C19 inhibitors) treatment with clopidogrel results in phenoconversion in genetically determined ultra-rapid phenotype to a poor metabolizer status indicated by loss of clopidogrel efficacy [35].

The beneficial side of phenoconversion interactions is that genetically determined phenotypes can be normalized by the addition of medications of opposite effects on metabolism. For example, resistance to nortriptyline (CYP2D6 substrate) due to abnormally rapid metabolism has been successfully reversed and normalized with the addition of paroxetine a (CYP2D6 inhibitor), which produces a recovery of nortriptyline therapeutic plasma levels [36].

Figure 3 presents different scenarios of phenoconversion interactions.

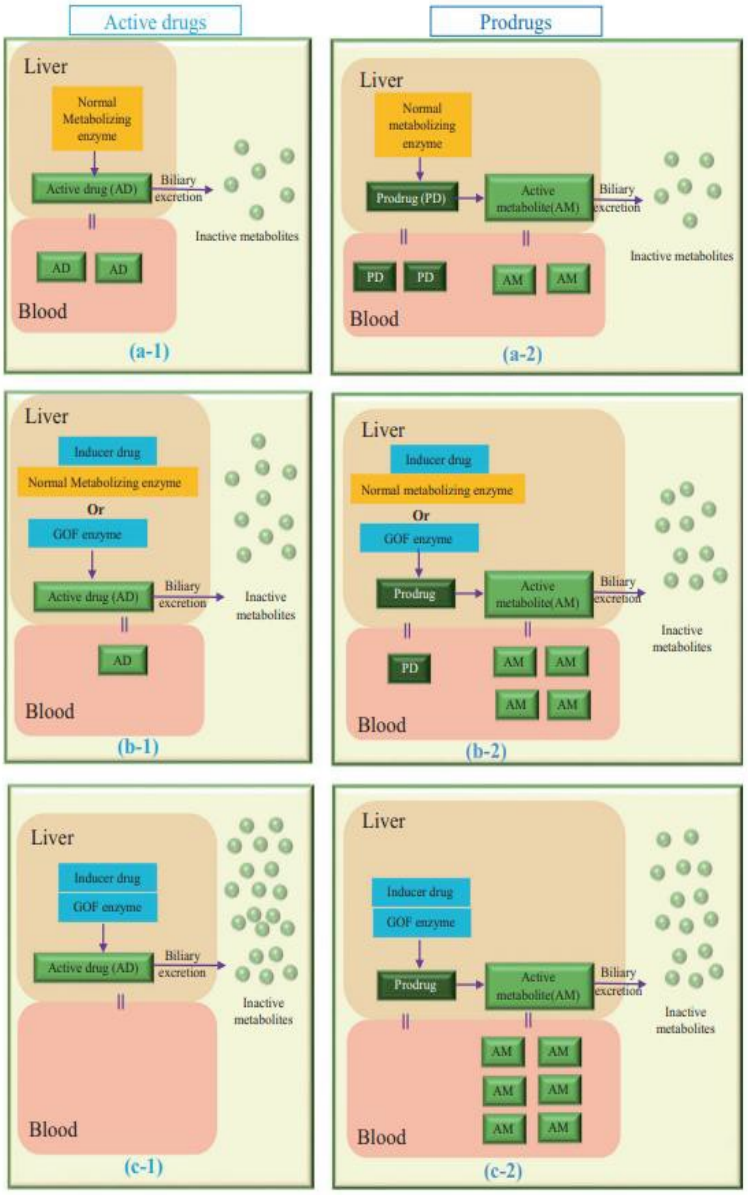


Fig. 2 The predicted active drug/active metabolites of prodrugs plasma levels and biliary excretion changes with out or with the presence of inducers or GOF variants or both on metabolizing enzymes. The predicted active drug/active metabolites of prodrugs plasma levels and biliary excretion changes without (**a-1/a-2**) or with the presence of inducers or GOF variants (**b-1/b-2**) or both (**c-1/c-2**) on metabolizing enzymes. **a-1/a-2** represent the normal scenario with no interacting drug or genetic variant. In **b-1/b-2** either an inducer drug or gain-of-function variant (GOF) in the metabolizing enzyme, results in increased metabolism to inactive metabolites, and decreased (**b-1**)/increased (**b-2**) active drug in the systemic circulation. In **c-1/c-2** the presence of inducer drug and the GOF genetic variant combine to produce greater decrease(**c-1**)/increase(**c-2**) in the systemic concentration of active drug

Drug–drug-transporters genes interactions (DDTGIs)

Drug transporters govern the movement of pharmaceutical compounds from and into different body tissues. The liver, kidney, blood–brain barrier (BBB), and intestine are the key sites of transporters that influence drug PK. In addition to summarizing the distribution and localization of transporters, Fig. 4 also classifies transporters into three categories according to the similarity of transport directions in different tissue types (the figure has been formulated with the aid of reference [37]). Drug–drug–gene interactions for transporters are less well studied than for metabolizing enzymes. For each subgroup, Drug Transporter-gene interaction (DTGI) studies will be utilized (if no direct DDTGI studies are available) to illustrate each mechanism for potential interaction. Similar to the drug metabolizing enzyme scenarios outlined above, we predict that these interactions may be intensified or reversed, via inhibitory/induction or phenoconversion pathways, with the co-administration of inhibitors or inducers.

Efflux transporters

Efflux transporters have been classified into two groups (group I and group II) according to the similarity in the transport directions.

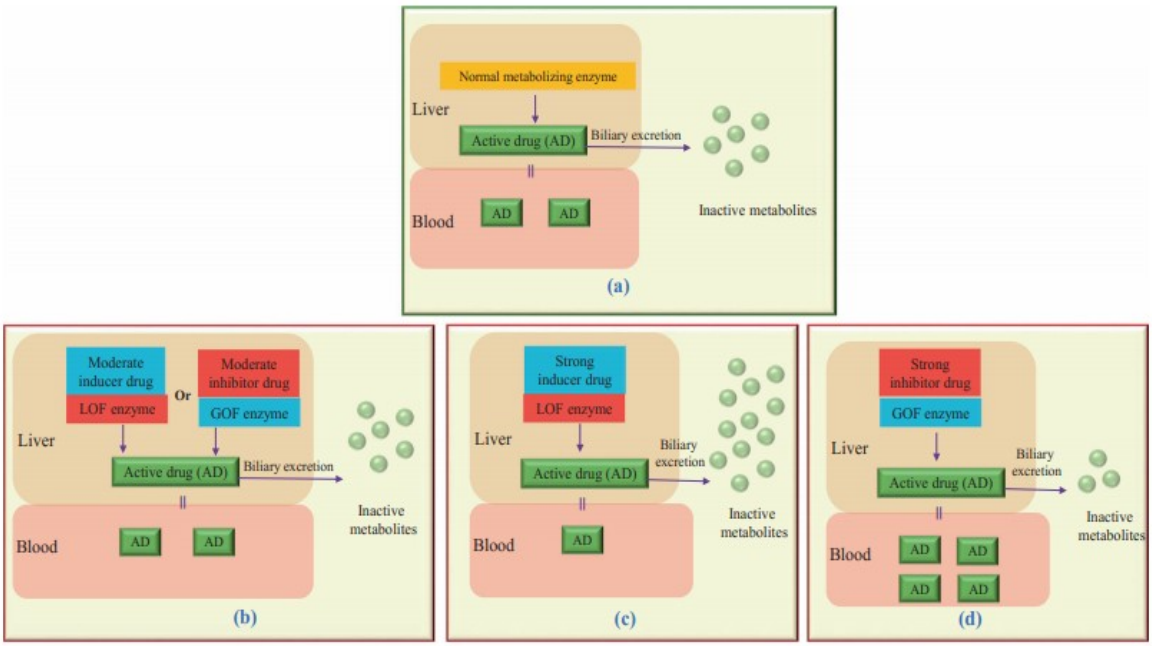


Fig. 3 Different scenarios of phenoconversion interactions where genetic effects may be reversed or shifted in the opposite direction. **a** Represents the normal scenario with no interacting drug or genetic variant. In **b** the effect of loss-of-function variant (LOF) or gain-of-function variant (GOF) is reversed with the presence of a moderate inducer drug or a moderate inhibitor drug respectively and results in a clinical outcome similar to the normal situation (**a**). In **c** the presence of a strong inducer drug has temporarily shifted a poor metabolism status into a rapid metabolism status and results in decreased active drug in the systemic circulation. In **d** the presence of a strong inhibitor drug has temporarily shifted a rapid metabolism status into a poor metabolism status and results in increased active drug in the systemic circulation

Group I P-glycoprotein 1 (P-gp, ABCB1), multidrug resistance-associated protein 2 (MRP2, ABCC2), and breast cancer resistant protein (BCRP, ABCG2) transporters are expressed in the intestine, liver, kidney, and BBB, sharing similar transport pathways. They efflux substrates back to intestinal lumen, facilitate hepatic and renal excretion (excluding BCRP), and work inversely in the BBB where they protect the brain from the entry of xenobiotics and return them back to systemic circulation. Blocking their function in the intestine, liver, or kidney is expected to elevate a substrate's systemic exposure (although opposite effects would be predicted if inhibiting transport across the BBB).

In this group, the most evidence for DDTGI comes from drugs altering ABCB1 (P-gp) transport and genetic variants in the gene encoding this transporter. For example, cyclosporine is an ABCB1 substrate. Diltiazem (a moderate ABCB1 inhibitor [38]) has been shown to increase cyclosporin trough concentrations in Chinese patients who carry the TT genotype (low P-gp activity) at rs1045642 (C>T) in ABCB1; yet no effect was seen in other ABCB1 genotypes (e.g., CC at rs1045642) [39]. Methadone is also a P-gp substrate, acting in the brain and effluxed across the BBB via P-gp. Patients with the TT genotype at rs1045642 and treated with quetiapine (ABCB1 inhibitor) experienced the lowest increase in methadone plasma levels compared with those with CT or CC genotypes (3% vs 23% vs 33% respectively) [40]. Low methadone plasma levels in this study would be explained by loss of the ABCB1 protective function in the BBB which results in increased intracerebral concentration of this central nervous system (CNS) drug. As a result of a similar DDTGI mechanism, the CNS drug granisetron was associated with increased efficacy in Japanese subjects (Supplementary Table 1) [41].

In some cases, it seems that adding strong inhibitors abolishes the effect of genotype. For example, no additional inhibitory effects were detected in carriers of different genotypes of the rs1045642 (C>T) ABCB1 variant who were either on dabigatran/rivaroxaban-clarithromycin combination or tacrolimus-itraconazole combination (ABCB1 substrates-ABCB1 strong inhibitors [38]) [42, 43].

ABCC2 and ABCG2 would be predicted to follow similar interaction scenarios as ABCB1, yet we were unable to find any studies that report DDTGIs for these transporters.

Group II Unlike group I transporters, there are no published studies describing DDTGIs for group II transporters. So here we report DGTIs to highlight the potential mechanisms whereby genes and drugs that alter these transporters may influence drug outcomes. MRP1(ABCC1), MRP3 (ABCC3), and MRP4(ABCC4) share the similar transport direction in the kidney and BBB as the Group I

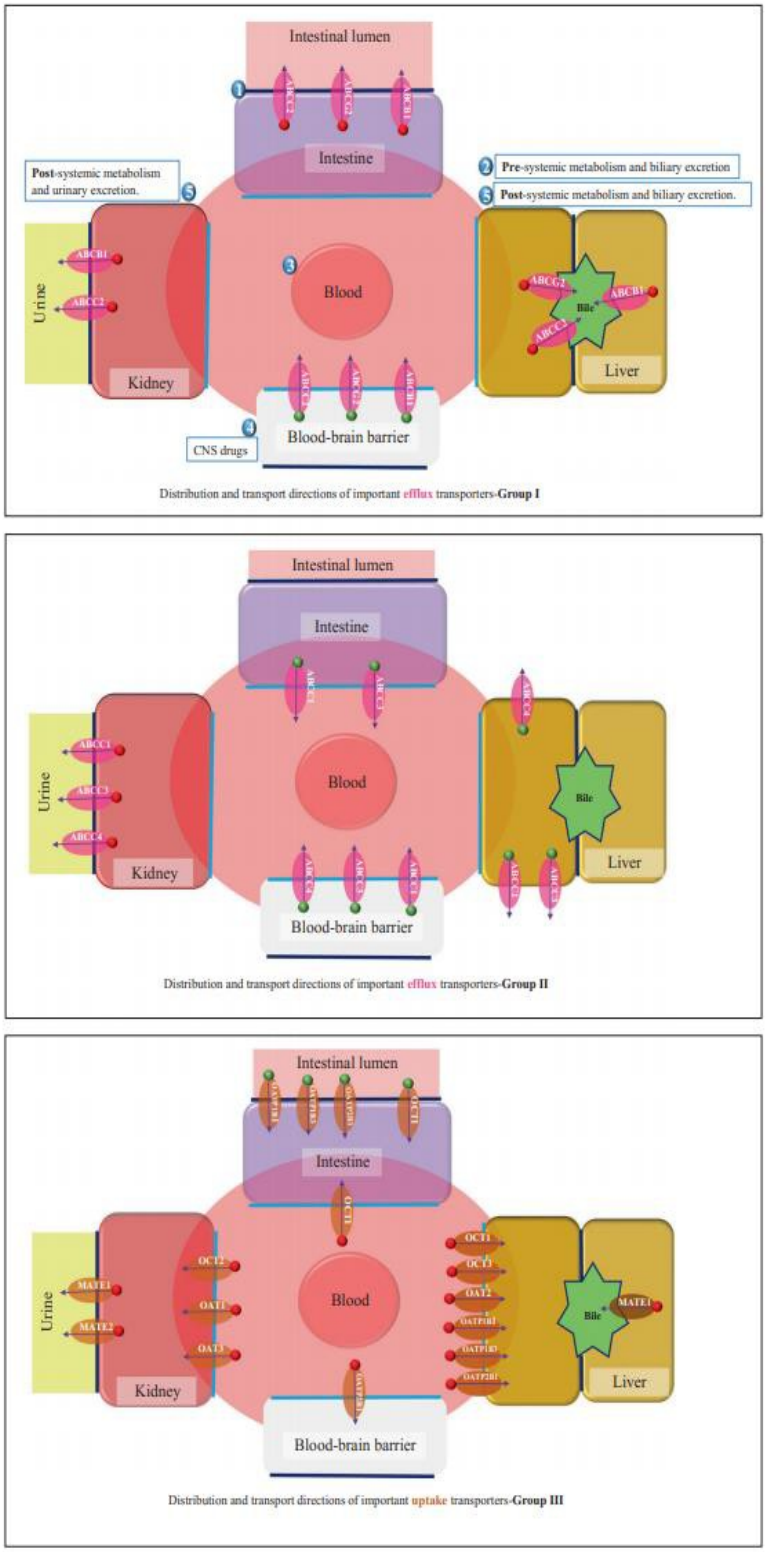


Fig. 4 Drug transporters as classified into three categories according to the similarity of the transport directions in different tissue types. Numbers from 1 to 5 = order of oral drug movement through different tissue types. Nonoral drug formulations bypass the effect of intestinal transporters. ●/● = increased/decreased substrate drug plasma level is predicted as a result of impairment of this transporter due to LOF variants or inhibitors. The reverse is predicted with GOF variants or inducers. The presence of the two factors (i.e. LOF variant + an inhibitor or GOF variant + an inducer) is predicted to double the clinical impact with neutralizing or shifting the clinical effect when the preparator drug and genetic effect are opposed (phenoconversion interactions). ——— = Apical membrane. ——— = Basolateral membrane

transporters. However, in the liver, they are expressed in the basolateral membrane working to pump drugs back into systemic circulation. MRP1, for example, transports the active metabolite of irinotecan (SN-38) out of hepatocytes into the blood contributing to the well-known side effect of irinotecan induced neutropenia [44]. The reduced function variant, rs17501331, in the *ABCC1* gene is associated with low incidence of neutropenia; the reverse effect was detected with the GOF variant rs6498588 in the same gene [45]. In some cases, increased activity of the MRP1 transporter can be advantageous, as seen with methotrexate hepatotoxicity where carriers of wild-type genotype of *ABCC1* rs246240 (A>G) variant are at higher risk for developing methotrexate toxicity compared with carriers of reduced function alleles [46]. Of note, MRP1 is also expressed in the myocardium protecting the heart from the entry of xenobiotics [47]. For example, the reduced transport associated with the rs45511401 (G>T) in ABCC1 increases the chance of developing cardiotoxicity due to intracellular accumulation of doxorubicin [48]. MRP1 and MRP3, in contrast to P-gp, MRP2 and BCRP, are expressed in the basolateral membrane of the intestine effluxing substrates into the portal circulation. As orally administered drugs are first exposed to intestinal transporters, any modification of their role might affect drug concentration in the other tissues (liver, kidney, or BBB). C.1037 C>T and c.1820G>A ABCC3 variants, for example, have low transport activity [49] suggesting their potential to diminish the bioavailability of oral MRP3 substrates irrespective of subsequent alteration in transport into other tissues, or subsequent metabolism.

Uptake Transporters (Group III)

In the liver, kidney, and BBB, all important uptake transporters (organic cation transporters (OCTs)1/2/3, organic anion-transporting polypeptide (OATP) 1B1/1B3/2B1, and multidrug and toxic compound extrusion proteins (MATE) 1/2), follow an identical main route for transporting their substrates: from systemic circulation into different tissues or urine/bile in case of MATEs. Consequently, reducing or increasing these transport capacities would result in increased or reduced systemic drug concentrations respectively. The reverse effects are seen with the uptake transporters expressed in the intestinal apical membrane such as OATPs and OCT1 since the transportation pathway is in the opposite direction.

In some circumstances, altering uptake transporter function can increase ADRs. For example, it has been observed that carriers of two OCT1 (*SLC22A1*) reduced function alleles who were treated with OCT1 inhibitors were over four times more likely to develop gastrointestinal side effects with metformin (an OCT1 substrate) treatment, which would be attributable to metformin accumulation in the intestinal lumen (assuming apical OCT1 localization) [50]. This finding was supported by a previous study [51]. At the level of renal uptake transporters, other DDTGIs have been reported in which carrying the mutant alleles and the co-administration of inhibitors was linked to increased metformin plasma levels/toxicity or reduced clearance (see Supplementary Table 1) [52, 53]. By contrast, reducing transport in some cases may reduce certain side effects. For instance, cisplatin (a OCT2 (*SLC22A2*) substrate) is both a nephrotoxic and an ototoxic agent. People carrying the rs316019 (C>A) OCT2 mutation were protected from these adverse reactions as the variant resulted in reduced transport of cisplatin into the kidney and the inner ear (cochlea) (where OCT2 is expressed as well) [54–56].

In many situations, the efficacy of a drug relies upon the ability of that drug to access certain tissues. Statins are taken up into the liver by OATP1B1(*SLCO1B1*) and this is crucial for their lipid lowering effect. Reducing this uptake pathway reduces statin efficacy and raises plasma concentrations, resulting in myopathy and, rarely, rhabdomyolysis. The rs4149056 (T>C) (SLCO1B1*15) variant has been widely studied, and in 23 studies [57–79], this variant has been persistently connected to increased statin plasma exposure, muscle aches, dose reduction, and/or treatment-resistant phenotypes. A number of other DDGIs have been described for the SLCO1B1 transporter. For example, although the increase in pravastatin (SLCO1B1 substrate) AUC after treatment with ritonavir (SLCO1B1 inhibitor) was not statistically significant (21% increase vs pravastatin alone) a large interaction was seen in those carrying the SLCO1B1*15 or *17 haplotypes, with a resulting 113% elevation in pravastatin AUC [80]. Other DDTGIs with the similar mechanism have also been published (see Supplementary Table 1) [81–83]. Interestingly, unlike the ritonavir example just outlined, in some situations reduced function variants do not show any significant PK change until after the addition of inhibitors. For example, patients with AG or AA genotypes at rs2289669(G>A) of the MATE1 transporter only had significantly lower metformin (MATE1 substrate) clearance compared with carriers of GG genotype after treatment with ranitidine (a MATE1 inhibitor) [84].

DDGIs and challenges in clinical practice

Metabolizing enzyme and transporter substrates, inducers, or inhibitors are not fully documented in many popular drug interaction databases, leaving physicians unaware of potentially important interactions. In addition, most of the resources commonly used by prescribers (e.g., Stockley's, Micromedex, Drug.com, RxList, or other drug interaction

checkers) do not consider genetic variation when classifying drug interactions into minor, moderate, or major classes. Genetic variation may markedly increase or ameliorate the severity of potential drug interactions and do need to be considered when considering real-world use of drugs.

This review has discussed the different mechanisms of interactions in their simplest forms with the assumption that the patient is free of transporter polymorphisms or inhibitors/inducers in the case of discussing DDMEGIs and vice versa with DDTGIs. However, in real-world clinical practice, achieving precisely tailored drug therapy requires a detailed examination of all mutations in the candidate enzyme or transporter genes with good awareness of the entire prescribed medications and possible pathways of interaction. Thus, the clinical scenario ranges from a relatively simple picture where the effect of genotype and interacting drug(s) can be approximated and treatment altered accordingly, to a far more complex scenario where physiologically based PK (PBPK) modeling may be helpful and where evaluation of large scale clinical data linked to genotypes is required to evaluate the clinical impact of multiple interacting drugs/multiple genotypes on drug outcomes.

Consider a relatively simple scenario: a patient with type 2 diabetes treated with metformin (has no effect on CYPs) who carries reduced function variants in *CYP2C9* (*2 or *3 variants) and who is started on gliclazide (CYP2C9/19 substrate). Reduced metabolism of gliclazide will result in increased efficacy [85] and increased risk of hypoglycemia [86]. The metformin use will not alter this DGI. However, if this patient were also treated with pioglitazone and/or atorvastatin (both are CYP2C9/19 inhibitors) they would be at potentially even greater risk of gliclazide-induced hypoglycemia and should be treated with a reduced dose of gliclazide. However, even for this simple scenario, such DDGI studies have not been reported; nor have dosing algorithms been developed to date for patients with CYP2C9 variants prescribed sulphonylureas and as such it is difficult to implement this into drug interaction calculators.

There are many more complex scenarios where, for example, a combination of both metabolizing enzyme and transporter LOF/GOF variants, as well as inhibitors/inducers are included. This kind of interaction may be only initially predictable when all their subinteractions result in the same clinical effect. For instance, reduced CYP3A4 and SLCO1B1 activities can both result in increased AUC of the substrate drug and a greater harm would be anticipated. Carriers of the TC genotype of SLCO1B1 rs4149056 (T>C) variant who are treated with amlodipine (CYP3A4 inhibitor) experienced a 90% increased simvastatin AUC compared with subjects not treated with amlodipine and wild-type for rs4149056 [87]. A similar scenario was reported with other two case reports (see Supplementary Table 1) [88, 89].

In other situations, subinteractions do not share a similar clinical effect. Here, predicting the overall clinical outcome is challenging. As an illustration, oral rosuvastatin is mainly eliminated via biliary excretion with a minor contribution of CYP2C9 to its metabolism [90]. This implies that its transporters (e.g., ABCC2, ABCG2, ABCC1, and SLCO1B1) are the core players in its elimination. The concomitant administration of verapamil (an ABCC1/2 inhibitor) and venlafaxine (an ABCG2 inducer) in those who have inherited CYP2C9*3 and/or SLCO1B1 rs4149056 (T>C) LOF variants results in unpredictable clinical consequences. CYP2C9, SLCO1B1, and ABCC2 impairment would boost rosuvastatin AUC, inducing ABCG2 would lower rosuvastatin AUC, and inhibition of ABCC1 could result in both increase or decrease in AUC (high AUC if the site of interaction is in the kidney and low AUC if it is in the intestine or liver). The exact estimation of the predicted net AUC following a certain DDGI relies on calculating the contribution of each metabolizing enzyme and transporter to the elimination process (i.e. degree of sensitivity of substrates), inhibition/induction potency of the perpetrator agent or the net effect of multiple inhibitors, inducers, or both, and the net percentage of reduction/elevation in the enzyme/s and/or transporter/s activity caused by a single or more SNPs. The outcome of such a hugely complex scenario is impossible to predict by the clinician, and requires a clinical support tool based upon a PK DDGI prediction algorithm. Most of the current work concentrates on generating either DD or DG interaction predictors rather than the combined effect of both drugs and variants. However, using PBPK models, one predictor tool (https://www.ddi-predictor.org/) has recently been successfully generated to estimate drug exposure and the recommended dose following the dual action of both the perpetrator drug and mutations in certain CYPs (CYP2D6, CYP2C9, and CYP2C19) [91]. Other PBPK models do attempt to incorporate genotype and drug–drug interactions, but these do not model transporter variants well and have yet to translate through into clinically useful tools [92].

An alternative method to evaluate the impact of DDGIs is via metabolizing enzymes and transporters endogenous biomarkers rather than plasma concentrations of substrate drugs. Multiple enzymes/transporters-related biomarkers have been identified [93]. For instance, it has been shown that the cholesterol, cortisone, and cortisol metabolites: 4β-hydroxycholestrol,6β-hydroxycortisone, and 6β-hydroxycortisole, respectively, which are catalyzed by CYP3A4 activity, are increased under the effect of inducers and decreased with inhibitors of CYP3A4. It was also recognized that bufotenine is a major metabolite resulting from the metabolizing activity of CYP2D6. With regard of transporters, several studies have observed the association between increased bilirubin plasma levels and

reduced hepatic OATP1B1/1B3 uptake function. The similar scenario was noted recently with the novel biomarkers coproporphyrins I and III (CPs I and III) where plasma CPs levels elevated with the inhibition of these transporters to a similar extent as with rosuvastatin. In DDGIs studies, endogenous biomarkers can be utilized to predict the effect of both genetic variants and inhibitors/inducers on the substrate drugs plasma levels.

It is worth noting that potential DDIs do not necessarily reflect actual interactions. It has been observed that clinically significant interactions are consistently lower than theoretically predictable interactions [94]. However, the authors noted that 20% of ADRs are linked with DDIs; most of them are serious with a high percentage of fatal cases. They also concluded that therapeutic failure secondary to DDIs, which is usually underestimated, represents a considerable part of total DDIs-related undesirable effects. The degree of clinical significance can be judged by observing other risk factors associated with a potential DDI such as polypharmacy and genetic variants. Polypharmacy is commonly seen with elderly and hospitalized patients making them the most vulnerable patient's subgroups to clinically significant interactions besides carriers of risky genetic variants. In addition, not all types of DDGIs are expected to be common. Induction and phenoconversion DDGIs are predicted to be seen with lower incidence compared with inhibitory DDGIs as the majority of perpetrator drugs are inhibitors rather than inducers and most of functional genetic variants are loss rather than GOF mutations.

The increasing availability of 'big data' linking health data and genomics has the potential to evaluate the real-world clinical impact of multiple drugs/multiple variant interactions. A number of data sets are now available or about to become available for study. In Scotland national prescribing and linked outcomes are available for the entire population enabling evaluation of real-world DDIs, and with an increasing bioresource (https://www.registerforshare.org) it should be possible to evaluate DDGIs in ~500 K people over the next few years. In addition, other resources such as UK biobank including genetic information on 500 K individuals (with primary care data available on 200 k during 2018) and other national bioresources (such as the Danish biorepository) and US bioresources linked to EHRs (EMERGE network) will enable the evaluation of n-way DDG interactions to identify clinically important interactions that can be incorporated into clinical decision support tools in the future.

Conclusion

Dozens of new pharmaceutical compounds enter the market each year and a considerable number of patients are prescribed multiple drugs that necessitate the utilization of drug interaction databases for better management. One of the major limitations of these drug interaction checkers is the omission of the genetic effect on drug interactions. This reflects both the lack of clinical studies that quantity potential DDGIs and the fact that genetic information is rarely available on patients at the point of prescribing. This review has illustrated, with some examples, various mechanisms by which DDGIs can occur at the level of metabolizing enzymes, drug transporters, or both (this has been summarized in Supplementary Table 1). We have also shown the different degrees of complexity clinicians may face in judging the predicted clinical outcome following a certain DDGI. The more factors that are included, the more challenging it becomes to evaluate the outcome. There is a need for PBPK models, clinical studies and real-world evaluation of drug outcomes linked to genetic information to develop clinical useful DDGI models, to reduce adverse DDIs and improve drug outcomes in the setting of increasing multi-morbidity and polypharmacy.

Compliance with ethical standards

Conflict of interest There is no conflict of interest to disclose that could be perceived to bias our work. This work is a part of an entire PhD project which is supported by the main author's sponsor for the purpose of completing his PhD degree. Authors and the sponsor have no conflict of interest to disclose.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Lazarou J, Pomernaz B, Corey P. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. Surv Anesthesiol. 1999;43:53–4.
2. Bouvy J, De Bruin M, Koopmanschap M. Epidemiology of adverse drug reactions in europe: a review of recent observational studies. Drug Saf 2015;38:437–53.
3. MiDatabank Adverse Drug Reaction (ADR) reporting – 2016 update, SPS—Specialist Pharmacy Service – The first stop for professional medicines advice. 2017. Accessed 6 Sep 2017. https://www.sps.nhs.uk/articles/midatabank-adverse-drug-reaction-adr-reporting-2016-update/.

4. Exploring the costs of unsafe care in the NHS. frontier-economics. 2014. Accessed 11 Sep 2017. http://www.frontier-economics.com/documents/2014/10/exploring-the-costs-of-unsafe-care-in-the-nhs-frontier-report-2-2-2-2.pdf.
5. Howard R, Avery A, Slavenburg S, Royal S, Pipe G, Lucassen P, et al. Which drugs cause preventable admissions to hospital? A systematic review. Br J Clin Pharmacol. 2007;63:136–47.
6. Rodrigues M, Oliveira C. Drug-drug interactions and adverse drug reactions in polypharmacy among older adults: an integrative review. Rev Lat Am de Enferm. 2016;24:e2800.
7. Verbeurgt P, Mamiya T, Oesterheld J. How common are drug and gene interactions? Prevalence in a sample of 1143 patients with CYP2C9, CYP2C19 and CYP2D6 genotyping. Pharmacogenomics 2014;15:655–65.
8. Bahar M, Setiawan D, Hak E, Wilffert B. Pharmacogenetics of drug–drug interaction and drug–drug–gene interaction: a systematic review on CYP2C9, CYP2C19 and CYP2D6. Pharmacogenomics 2017;18:701–39.
9. Andersson M, Eliasson E, Lindh J. A clinically significant interaction between warfarin and simvastatin is unique to carriers of theCYP2C9*3allele. Pharmacogenomics 2012;13:757–62.
10. Malhi H. Warfarin and celecoxib interaction in the setting of cytochrome P450 (CYP2C9) polymorphism with bleeding complication. Postgrad Med J. 2004;80:107–9.
11. Uno T, Shimizu M, Yasui-Furukori N, Sugawara K, Tateishi T. Different effects of fluvoxamine on rabeprazole pharmacokinetics in relation to CYP2C19 genotype status. Br J Clin Pharmacol. 2006;61:309–14.
12. Yasui-Furukori N, Saito M, Uno T, Takahata T, Sugawara K, Tateishi T. Effects of fluvoxamine on lansoprazole pharmacokinetics in relation to CYP2C19 genotypes. J Clin Pharmacol. 2004;44:1223–9.
13. Cho J, Yu K, Jang I, Yang B, Shin S, Yim D. Omeprazole hydroxylation is inhibited by a single dose of moclobemide in homozygotic EM genotype for CYP2C19. Br J Clin Pharmacol. 2002;53:393–7.
14. Hamelin B. Significant interaction between the nonprescription antihistamine diphenhydramine and the CYP2D6 substrate metoprolol in healthy men with high or low CYP2D6 activity. Clin Pharmacol Therapeutics. 2000;67:466–77.
15. Damy T, Pousset F, Caplain H, Hulot J, Lechat P. Pharmacokinetic and pharmacodynamic interactions between metoprolol and dronedarone in extensive and poor CYP2D6 metabolizers healthy subjects. Fundam Clin Pharmacol. 2004;18:113–23.
16. Zhu L, Brüggemann R, Uy J, Colbers A, Hruska M, Chung E, et al. CYP2C19 Genotype-Dependent Pharmacokinetic Drug Interaction Between Voriconazole and Ritonavir-Boosted Atazanavir in Healthy Subjects. J Clin Pharm. 2017;57:235–46.
17. Imamura C, Furihata K, Okamoto S, Tanigawara Y. Impact of cytochrome P450 2C19 polymorphisms on the pharmacokinetics of tacrolimus when coadministered with voriconazole. J Clin Pharmacol. 2015;56:408–13.
18. Shi H, Yan J, Zhu W, Yang G, Tan Z, Wu W, et al. Effects of erythromycin on voriconazole pharmacokinetics and association with CYP2C19 polymorphism. Eur J Clin Pharmacol. 2010;66:1131–6.
19. Saito M, Yasui-Furukori N, Uno T, Takahata T, Sugawara K, Munakata A, et al. Effects of clarithromycin on lansoprazole pharmacokinetics between CYP2C19 genotypes. Br J Clin Pharmacol. 2005;59:302–9.
20. Sangkuhl K, Klein T, Altman R. Clopidogrel pathway. Pharmacogenetics Genomics. 2010;20:463–5.
21. Furuta T, Iwaki T, Umemura K. Influences of different proton pump inhibitors on the anti-platelet function of clopidogrel in relation to CYP2C19 genotypes. Br J Clin Pharmacol. 2010;70:383–92.
22. Harmze A, Van Werkum J, Souverein P, Breet N, Bouman H, Hackeng C, et al. Combined influence of proton-pump inhibitors, calcium-channel blockers and CYP2C19*2 on on-treatment platelet reactivity and on the occurrence of atherothrombotic events after percutaneous coronary intervention. J Thrombosis Haemost. 2011;9:1892–901.
23. Malingré M, Godschalk P, Klein S. A case report of voriconazole therapy failure in a homozygous ultrarapid CYP2C19*17/*17 patient comedicated with carbamazepine. Br J Clin Pharmacol. 2012;74:205–6.
24. Carlquist J, Knight S, Horne B, Huntinghouse J, Rollo J, Muhlestein J, et al. Cardiovascular risk among patients on clopidogrel anti-platelet therapy after placement of drug-eluting stents is modified by genetic variants in both the CYP2C19 and ABCB1 genes. Thrombosis Haemost. 2013;109:744–54.
25. Chan M, Tan K, Tan H, Huan P, Li B, Phua Q, et al. CYP2C19 and PON1 polymorphisms regulating clopidogrel bioactivation in Chinese, Malay and Indian subjects. Pharmacogenomics 2012;13:533–42.
26. Harmsze A, van Werkum J, Hackeng C, Ruven H, Kelder J, Bouman H, et al. The influence of CYP2C19*2 and *17 on on-treatment platelet reactivity and bleeding events in patients undergoing elective coronary stenting. Pharmacogenetics Genomics. 2012;22:169–75.
27. Cayla G, Hulot J, O'Connor S, Pathak A, Scott S, Gruel Y, et al. Clinical, angiographic, and genetic factors associated with early coronary stent thrombosis. JAMA 2011;306:1765–74.
28. Gurbel P, Shuldiner A, Bliden K, Ryan K, Pakyz R, Tantry U. The relation between CYP2C19 genotype and phenotype in stented patients on maintenance dual antiplatelet therapy. Am Heart J. 2011;161:598–604.
29. Wallentin L, James S, Storey R, Armstrong M, Barratt B, Horrow J, et al. Effect of CYP2C19 and ABCB1 single nucleotide polymorphisms on outcomes of treatment with ticagrelor versus clopidogrel for acute coronary syndromes: a genetic substudy of the PLATO trial. Lancet 2010;376:1320–8.
30. Sibbing D, Gebhard D, Koch W, Braun S, Stegherr J, Morath T, et al. Isolated and interactive impact of common CYP2C19 genetic variants on the antiplatelet effect of chronic clopidogrel therapy. J Thrombosis Haemost. 2010;8:1685–93.
31. Gladding P, Webster M, Zeng I, Farrell H, Stewart J, Ruygrok P, et al. The pharmacogenomics and pharmacodynamics of clopidogrel response: an analysis from the PRINC (Plavix Response in Coronary Intervention) trial. Heart, Lung Circulation. 2008;17:S82.
32. Dai Z. Relationship between cytochrome P450 2C19*17 genotype distribution platelet aggregation and bleeding risk in patients with blood stasis syndrome of coronary artery disease treated with clopidogrel. J Chin Integr Med. 2012;10:647–54.
33. Tiroch K, Sibbing D, Koch W, Roosen-Runge T, Mehilli J, Schömig A, et al. Protective effect of the CYP2C19 *17 polymorphism with increased activation of clopidogrel on cardiovascular events. Am Heart J. 2010;160:506–12.
34. Vormfelde S, Brockmöller J, Bauer S, Herchenhein P, Kuon J, Meineke I, et al. Relative Impact of Genotype and Enzyme Induction on the Metabolic Capacity of CYP2C9 in Healthy Volunteers. Clin Pharmacol Therapeutics. 2009;86:54–61.
35. Depta J, Lenzini P, Lanfear D, Wang T, Spertus J, Bach R, et al. Clinical outcomes associated with proton pump inhibitor use among clopidogrel-treated patients within CYP2C19 genotype groups following acute myocardial infarction. Pharmacogenomics J. 2014;15:20–5.
36. Laine K, Tybring G, Hartter S, Andersson K, Svensson J, Widen J, et al. Inhibition of cytochrome P4502D6 activity with paroxetine normalizes the ultrarapid metabolizer phenotype as measured

by nortriptyline pharmacokinetics and the debrisoquin test. Clin Pharmacol Therapeutics. 2001;70:327–35.
37. Nicholls G, Youdim K. Drug transporters: volume 1: role and importance in adme and drug development, p.16. 54th edn. The Royal Society of Chmeistry (RSC) 2016; 2016. Accessed 9 Sep 2017. http://pubs.rsc.org/en/content/ebook/978-1-78262-069-3.
38. Wessler J, Grip L, Mendell J, Giugliano R. The P-glycoprotein transport system and cardiovascular drugs. J Am Coll Cardiol. 2013;61:2495–502.
39. Wang Y, Li J, Wang X, Zhang Y, Wang C, Huang M. Diltiazem augments the influence of MDR1 genotype status on cyclosporine concentration in Chinese patients with renal transplantation. Acta Pharmacologica Sin. 2015;36:855–62.
40. Uehlinger C, Crettol S, Chassot P, Brocard M, Koeb L, Brawand-Amey M, et al. Increased (R)-methadone plasma concentrations by quetiapine in cytochrome P450s and ABCB1 genotyped patients. J Clin Psychopharmacol. 2007;27:273–8.
41. Tsuji D, Kim Y, Nakamichi H, Daimon T, Suwa K, Iwabe Y, et al. Association of ABCB1 polymorphisms with the antiemetic efficacy of granisetron plus dexamethasone in breast cancer patients. Drug Metab Pharmacokinetics. 2013;28:299–304.
42. Gouin-Thibault I, Delavenne X, Blanchard A, Siguret V, Salem J, Narjoz C, et al. Interindividual variability in dabigatran and rivaroxaban exposure: contribution ofABCB1genetic polymorphisms and interaction with clarithromycin. J Thrombosis Haemost. 2017;15:273–83.
43. Togashi M, Niioka T, Komatsuda A, Nara M, Okuyama S, Omokawa A, et al. Effect of CYP3A5 and ABCB1 polymorphisms on the interaction between tacrolimus and itraconazole in patients with connective tissue disease. Eur J Clin Pharmacol. 2015;71:1091–7.
44. Lowenberg D, Thorn C, Whirl-Carrillo M, Ramirez J, Gong L, Marsh S et al. Irinotecan Pathway, Pharmacokinetics. PharmGKB. Accessed 9 Oct 2017. https://www.pharmgkb.org/pathway/PA2001.
45. Li M, Seiser E, Baldwin R, Ramirez J, Ratain M, Innocenti F, et al. ABC transporter polymorphisms are associated with irinotecan pharmacokinetics and neutropenia. Pharmacogenomics J. 2016. https://doi.org/10.1038/tpj.2016.75.
46. Warren R, Smith R, Campalani E, Eyre S, Smith C, Barker J, et al. Genetic variation in efflux transporters influences outcome to methotrexate therapy in patients with psoriasis. J Investig Dermatol. 2008;128:1925–9.
47. Jungsuwadee P, Nithipongvanitch R, Chen Y, Oberley T, Butterfield D,St, Clair D, et al. Mrp1 Localization and Function in Cardiac Mitochondria after Doxorubicin. Mol Pharmacol 2009;75:1117–26.
48. Wojnowski L. NAD(P)H oxidase and multidrug resistance protein genetic polymorphisms are associated with doxorubicin-induced cardiotoxicity. Circulation 2005;112:3754–62.
49. Kobayashi K, Ito K, Takada T, Sugiyama Y, Suzuki H. Functional analysis of nonsynonymous single nucleotide polymorphism type ATP-binding cassette transmembrane transporter subfamily C member 3. Pharmacogenetics Genomics. 2008;18:823–33.
50. Dujic T, Zhou K, Donnelly L, Tavendale R, Palmer C, Pearson E. Association of organic cation transporter 1 with intolerance to metformin in type 2 diabetes: a GoDARTS study. Diabetes 2014;64:1786–93.
51. Ahlin G, Chen L, Lazorova L, Chen Y, Ianculescu A, Davis R, et al. Genotype-dependent effects of inhibitors of the organic cation transporter, OCT1: predictions of metformin interactions. Pharmacogenomics J. 2010;11:400–11.
52. Grün B, Kiessling M, Burhenne J, Riedel K, Weiss J, Rauch G, et al. Trimethoprim-metformin interaction and its genetic modulation by OCT2 and MATE1 transporters. Br J Clin Pharmacol. 2013;76:787–96.
53. Wang Z, Yin O, Tomlinson B, Chow M. OCT2 polymorphisms and in-vivo renal functional consequence: studies with metformin and cimetidine. Pharmacogenetics Genomics. 2008;18:637–45.
54. Filipski K, Mathijssen R, Mikkelsen T, Schinkel A, Sparreboom A. Contribution of organic cation transporter 2 (OCT2) to cisplatin-induced nephrotoxicity. Clin Pharmacol Therapeutics. 2009;86:396–402.
55. Spracklen T, Vorster A, Ramma L, Dalvie S, Ramesar R. Promoter region variation in NFE2L2 influences susceptibility to ototoxicity in patients exposed to high cumulative doses of cisplatin. Pharmacogenomics J. 2016. https://doi.org/10.1038/tpj.2016.52.
56. Lanvers-Kaminsky C, Sprowl J, Malath I, Deuster D, Eveslage M, Schlatter E, et al. Human OCT2 variant c.808G>T confers protection effect against cisplatin-induced ototoxicity. Pharmacogenomics 2015;16:323–32.
57. Choi H, Bae K, Cho S, Ghim J, Choe S, Jung J, et al. Impact of CYP2D6, CYP3A5, CYP2C19, CYP2A6, SLCO1B1, ABCB1, and ABCG2 gene polymorphisms on the pharmacokinetics of simvastatin and simvastatin acid. Pharmacogenet Genomics 2013;25:595–608.
58. Luzum J, Theusch E, Taylor K, Wang A, Sadee W, Binkley P, et al. Individual and combined associations of genetic variants in CYP3A4, CYP3A5, and SLCO1B1 with simvastatin and simvastatin acid plasma concentrations. J Cardiovascular Pharmacol. 2015;66:80–5.
59. Mirošević Skvrce N, Macolić Šarinić V, Šimić I, Ganoci L, Muačević Katanec D, Božina N. ABCG2 gene polymorphisms as risk factors for atorvastatin adverse reactions: a case-control study. Pharmacogenomics 2015;16:803–15.
60. Tsamandouras N, Dickinson G, Guo Y, Hall S, Rostami-Hodjegan A, Galetin A, et al. Development and application of a mechanistic pharmacokinetic model for simvastatin and its active metabolite simvastatin acid using an integrated population PBPK approach. Pharm Res 2014;32:1864–83.
61. Birmingham B, Bujac S, Elsby R, Azumaya C, Wei C, Chen Y, et al. Impact of ABCG2 and SLCO1B1 polymorphisms on pharmacokinetics of rosuvastatin, atorvastatin and simvastatin acid in Caucasian and Asian subjects: a class effect? Eur J Clin Pharmacol. 2015;71:341–55.
62. Tsamandouras N, Dickinson G, Guo Y, Hall S, Rostami-Hodjegan A, Galetin A, et al. Identification of the effect of multiple polymorphisms on the pharmacokinetics of simvastatin and simvastatin acid using a population-modeling approach. Clin Pharmacol Ther. 2014;96:90–100.
63. Ferrari M, Guasti L, Maresca A, Mirabile M, Contini S, Grandi A, et al. Association between statin-induced creatine kinase elevation and genetic polymorphisms in SLCO1B1, ABCB1 and ABCG2. Eur J Clin Pharmacol. 2014;70:539–47.
64. De Keyser C, Peters B, Becker M, Visser L, Uitterlinden A, Klungel O, et al. The SLCO1B1 c.521T>C polymorphism is associated with dose decrease or switching during statin therapy in the Rotterdam Study. Pharmacogenetics Genomics. 2014;24:43–51.
65. DeGorter M, Tirona R, Schwarz U, Choi Y, Dresser G, Suskin N, et al. Clinical and pharmacogenetic predictors of circulating atorvastatin and rosuvastatin concentrations in routine clinical care. Circulation. 2013;6:400–8.
66. Carr D, O'Meara H, Jorgensen A, Campbell J, Hobbs M, McCann G, et al. SLCO1B1 genetic variant associated with statin-induced myopathy: a proof-of-concept study using the clinical practice research datalink. Clin Pharmacol Ther. 2013;94:695–701.
67. Ulvestad M, Skottheim I, Jakobsen G, Bremer S, Molden E, Åsberg A, et al. Impact of OATP1B1, MDR1, and CYP3A4

expression in liver and intestine on interpatient pharmacokinetic variability of atorvastatin in obese subjects. Clin Pharmacol Therap. 2012;93:275–82.
68. Marciante K, Durda J, Heckbert S, Lumley T, Rice K, McKnight B, et al. Cerivastatin, genetic variants, and the risk of rhabdomyolysis. Pharmacogenet Genomics 2011;21:280–8.
69. Donnelly L, Doney A, Tavendale R, Lang C, Pearson E, Colhoun H, et al. Common nonsynonymous substitutions in SLCO1B1 predispose to statin intolerance in routinely treated individuals with type 2 diabetes: A Go-DARTS Study. Clin Pharmacol Therapeutics. 2010;89:210–6.
70. Voora D, Shah S, Spasojevic I, Ali S, Reed C, Salisbury B, et al. The SLCO1B1*5 genetic variant is associated with statin-induced side effects. J Am Coll Cardiol. 2009;54:1609–16.
71. Link E, Parish S, Armitage J, Bowman L, Heath S, Matsuda F, et al. SLCO1B1 variants and statin-induced myopathy-a genomewide study. N. Engl J Med. 2008;359:789–99.
72. Ho R, Choi L, Lee W, Mayo G, Schwarz U, Tirona R, et al. Effect of drug transporter genotypes on pravastatin disposition in European- and African-American participants. Pharmacogenetics Genomics. 2007;17:647–56.
73. Pasanen M, Fredrikson H, Neuvonen P, Niemi M. Different effects of SLCO1B1 polymorphism on the pharmacokinetics of atorvastatin and rosuvastatin. Clin Pharmacol Ther. 2007;82:726–33.
74. Niemi M, Pasanen M, Neuvonen P. SLCO1B1 polymorphism and sex affect the pharmacokinetics of pravastatin but not fluvastatin. Clin Pharmacol Ther. 2006;80:356–66.
75. Igel M, Arnold K, Niemi M, Hofmann U, Schw ABM, Lutjohann D, et al. Impact of the SLCO1B1 polymorphism on the pharmacokinetics and lipid-lowering efficacy of multiple-dose pravastatin. Clin Pharmacol Therapeutics. 2006;79:419–26.
76. Pasanen M, Neuvonen M, Neuvonen P, Niemi M. SLCO1B1 polymorphism markedly affects the pharmacokinetics of simvastatin acid. Pharmacogenetics Genomics. 2006;16:873–9.
77. Thompson J, Man M, Johnson K, Wood L, Lira M, Lloyd D, et al. An association study of 43 SNPs in 16 candidate genes with atorvastatin response. Pharmacogenomics J. 2005;5:352–8.
78. Mwinyi J. Evidence for inverse effects of OATP-C (SLC21A6) *5 and *1b haplotypes on pravastatin kinetics. Clin Pharmacol Therapeutics. 2004;75:415–21.
79. Nishizato Y. Polymorphisms of OATP-C (SLC21A6) and OAT3 (SLC22A8) genes: consequences for pravastatin pharmacokinetics. Clin Pharmacol Ther. 2003;73:554–65.
80. Aquilante C, Kiser J, Anderson P, Christians U, Kosmiski L, Daily E, et al. Influence of SLCO1B1 polymorphisms on the drug-drug interaction between darunavir/ritonavir and pravastatin. J Clin Pharmacol. 2012;52:1725–38.
81. Suwannakul S, Ieiri I, Kimura M, Kawabata K, Kusuhara H, Hirota T, et al. Pharmacokinetic interaction between pravastatin and olmesartan in relation to SLCO1B1 polymorphism. J Hum Genet. 2008;53:899–904.
82. Kalliokoski A, Backman J, Kurkinen K, Neuvonen P, Niemi M. Effects of gemfibrozil and atorvastatin on the pharmacokinetics of repaglinide in relation to SLCO1B1 polymorphism. Clin Pharmacol Ther. 2008;84:488–96.
83. He Y, Zhang W, Chen Y, Guo D, Tu J, Xu L, et al. Rifampicin alters atorvastatin plasma concentration on the basis of SLCO1B1 521T>C polymorphism. Clin Chim Acta. 2009;405:49–52.
84. Cho S, Chung J. The MATE1 rs2289669 polymorphism affects the renal clearance of metformin following ranitidine treatment. Int J Clin Pharmacol Ther. 2016;54:253–62.
85. Zhou K, Donnelly L, Burch L, Tavendale R, Doney A, Leese G, et al. Loss-of-Function CYP2C9 variants improve therapeutic response to sulfonylureas in type 2 diabetes: A Go-DARTS Study. Clin Pharmacol Ther. 2009;87:52–6.
86. Ragia G, Petridis I, Tavridou A, Christakidis D, Manolopoulos V. Presence of CYP2C9*3 allele increases risk for hypoglycemia in Type 2 diabetic patients treated with sulfonylureas. Pharmacogenomics 2009;10:1781–7.
87. Jiang F, Choi J, Lee J, Ryu S, Park Z, Lee J, et al. The influences of SLCO1B1 and ABCB1 genotypes on the pharmacokinetics of simvastatin, in relation to CYP3A4 inhibition. Pharmacogenomics 2017;18:459–69.
88. Meyer zu Schwabedissen H, Siegmund W, Kroemer H, Rollnik J. Creatine kinase elevation caused by a combination of fluvastatin and telmisartan in a patient heterozygous for the CYP2C9*3 and ABCC2 -24C>T variants: a case report. BMC Res Notes. 2014;7:688.
89. Marusic S, Lisicic A, Horvatic I, Bacic-Vrca V, Bozina N. Atorvastatin-related rhabdomyolysis and acute renal failure in a genetically predisposed patient with potential drug–drug interaction. Int J Clin Pharm. 2012;34:825–7.
90. Rosuvastatin - DrugBank. Drugbank.ca. Accessed 9 Oct 2017. https://www.drugbank.ca/drugs/DB01098.
91. Tod M, Nkoud-Mongo C, Gueyffier F. Impact of genetic polymorphism on drug-drug interactions mediated by cytochromes: a general approach. AAPS J. 2013;15:1242–52.
92. Zhuang X, Lu C. PBPK modeling and simulation in drug research and development. Acta Pharmaceutica Sin B 2016;6:430–40.
93. Mariappan T, Shen H, Marathe P. Endogenous biomarkers to assess drug-drug interactions by drug transporters and enzymes. Curr Drug Metab. 2017;18:757–68.
94. Magro L, Moretti U, Leone R. Epidemiology and characteristics of adverse drug reactions caused by drug–drug interactions. Expert Opin Drug Saf. 2011;11:83–94.

This page was left intentionally

Supplementary material 2:

The final 162 selected SNPs with their MAF and HWE results from all cohorts.
*SNP groups: 1= 26 important SNPs, 2 = 76 independent SNPs, and 3 = 60 randomly selected SNPs among groups of correlated SNPs.

									GoDARTs		GS		GoSHARE		Combined		UKBB	
No.	rsid	Position_build_37	SNP group *	Major allele	Minor allele	* allele	Gene	Chromosome	MAF	HWE (p-value)	MAF	HWE (p-value)	MAF	HWE (p-value)	MAF	HWE (p-value)	MAF	HWE (p-value)
1	rs1045642	87138645	1	A	G	*2/*13	*ABCB1*	7	0.46361993	0.301747799	0.458067093	0.714597532	0.464754953	0.203217862	0.460458956	0.440908535	0.45625	0.06113509
2	rs1128503	87179601	1	G	A	NA	*ABCB1*	7	0.44226689	0.078822311	0.452500998	0.682991294	0.449322211	0.151744141	0.449424367	0.257672978	0.4371464	0.06359866
3	rs2032582	87160618	1	C	A	NA	*ABCB1*	7	0.460066808	0.456101958	0.46083766	0.745763601	0.456517205	0.02989111	0.460092725	0.17082713	0.4561078	0.3049629
4	rs1186746	87133993	2	T	C	NA	*ABCB1*	7	0.259061513	0.343387799	0.127720647	0.176156577	0.121272158	0.782606731	0.124939978	0.457945631	0.1137491	0.07089484
5	rs12720066	87169702	2	A	C	NA	*ABCB1*	7	0.056235717	0.96149662	0.259210264	0.781268221	0.253180396	0.480632005	0.258226228	0.277776174	0.2720129	0.459068
6	rs2032588	87179443	2	G	A	NA	*ABCB1*	7	0.068757674	0.308036192	0.057657748	0.609020715	0.064859228	0.685035282	0.060099818	0.129446994	0.06216873	*< 1*10-8*
7	rs3842	87133366	2	T	C	NA	*ABCB1*	7	0.155827322	0.09014297	0.056908946	0.727020808	0.064025026	0.774576965	0.059252738	0.206786659	0.06076218	0.003384094
8	rs4728709	87233602	2	G	A	NA	*ABCB1*	7	0.061397633	0.490482721	0.249575679	0.810723151	0.257142857	0.78232084	0.249698395	0.642264407	0.2510837	0.09216039
9	rs9282564	87229440	2	T	C	NA	*ABCB1*	7	0.097764322	0.03821894	0.216728235	0.781621444	0.215015641	0.591835595	0.214976562	0.915007841	0.214707	0.7667071
10	rs1186745	87133947	3	C	A	NA	*ABCB1*	7	0.128333085	0.902763069	0.146440695	0.423916518	0.15265902	0.970630609	0.145362293	0.871344139	0.1492706	0.6086663
11	rs2235013	87178626	3	C	T	NA	*ABCB1*	7	0.47054229	0.488204875	0.117462061	0.936549577	0.118456726	0.799306212	0.118842693	0.317287328	0.1223943	0.9279406
12	rs4148739	87161049	3	T	C	NA	*ABCB1*	7	0.134887872	0.674554603	0.047898363	0.700544836	0.051616267	0.497310605	0.051523771	0.976466428	0.05285953	0.1522764
13	rs152023	16085236	2	T	C	NA	*ABCC1*	16	0.306932656	0.728744821	0.117561901	0.982463359	0.118039625	0.849683212	0.119032775	0.341922155	0.122257	0.816522
14	rs16967126	16128418	2	T	C	NA	*ABCC1*	16	0.08373393	0.755782298	0.127346246	0.149035713	0.131386861	0.052379586	0.127272997	0.231492416	0.1315234	0.849175
15	rs17287570	16155103	2	A	C	NA	*ABCC1*	16	0.191023567	0.904572321	0.248028155	0.725571509	0.238894682	0.15425162	0.24710734	0.589360208	0.2464634	0.3060833
16	rs17501331	16089441	2	A	G	NA	*ABCC1*	16	0.092577432	0.114325292	0.064621605	0.981327652	0.064650678	0.042635067	0.064913211	0.701081242	0.06276541	0.5430438
17	rs2074087	16184232	2	G	C	NA	*ABCC1*	16	0.13935991	0.045708425	0.123901757	0.47277556	0.125860271	0.956578786	0.125876185	0.088155977	0.1352321	0.9720862
18	rs212090	16236004	2	A	T	NA	*ABCC1*	16	0.448242521	0.50870511	0.210163738	0.65883872	0.215745568	0.707421751	0.21425569	0.565516127	0.2219529	*< 1*10-8*
19	rs215095	16060394	2	A	G	NA	*ABCC1*	16	0.116290811	0.916192172	0.456594449	0.844980188	0.449009385	0.402512745	0.452368073	0.286879544	0.4518098	*< 1*10-8*
20	rs2889517	16181956	2	C	T	NA	*ABCC1*	16	0.280029435	0.528310133	0.329248203	0.062754193	0.33096976	0.877420561	0.327639333	0.579722081	0.325659	*< 1*10-8*
21	rs35621	16168608	2	C	T	NA	*ABCC1*	16	0.132808135	0.969712542	0.100938498	0.660367074	0.102294056	0.709054389	0.099952333	0.845755591	0.09608641	6.54E-08
22	rs3743527	16235681	2	C	T	NA	*ABCC1*	16	0.196559504	0.682366402	0.200554113	0.747444158	0.202815433	0.76512132	0.202875112	0.535614905	0.2075959	*< 1*10-8*
23	rs8187843	16101875	2	G	A	NA	*ABCC1*	16	0.081941725	0.545553962	0.083191893	0.584717368	0.089259645	0.192375021	0.083814169	0.009833538	0.08851212	0.5950531
24	rs11861115	16199670	3	G	A	NA	*ABCC1*	16	0.24635405	0.007851622	0.064771366	0.946349511	0.062252346	0.782792169	0.064879513	0.553938188	0.07002605	0.8726684
25	rs1967120	16108894_16108895	3	A	G	NA	*ABCC1*	16	0.291950811	0.533475283	0.39057508	0.345814587	0.394786236	0.387759192	0.388478706	0.1419192	0.3891436	0.4411858
26	rs212082	16227147	3	A	G	NA	*ABCC1*	16	0.205325112	0.818787982	0.162365216	0.39204648	0.160896767	0.794246535	0.161047051	0.760200016	0.1500915	0.3749078
27	rs215066	16079117	3	G	A	NA	*ABCC1*	16	0.056142144	0.952080151	0.259784345	0.100673993	0.267049009	0.78806089	0.260717926	0.285316079	0.2581483	1.76E-07
28	rs35596	16152940	3	T	C	NA	*ABCC1*	16	0.2126378	0.13573572	0.064496805	0.248236476	0.063399374	0.759481522	0.061636801	0.413787494	0.06176809	0.1723977
29	rs45511401	16173232	3	G	T	NA	*ABCC1*	16	0.055733945	0.16282317	0.065869609	0.072949371	0.064754953	0.17887041	0.066471845	0.145063054	0.06066049	0.5065297
30	rs4781712	16103232	3	G	A	NA	*ABCC1*	16	0.448702101	0.48508198	0.15634984	0.596696035	0.149426486	0.451358371	0.155120616	0.465953225	0.1399296	0.08815652
31	rs8058040	16107712	3	A	G	NA	*ABCC1*	16	0.168973987	0.264512895	0.060203674	0.907936158	0.060896767	0.111014416	0.060569648	0.984877753	0.0644072	0.5916309
32	rs2273697	101563815	2	G	A	NA	*ABCC2*	10	0.197784086	0.440369379	0.102535942	0.88161535	0.098540146	0.856953662	0.100600304	0.230125499	0.109195	0.3056447
33	rs717620	101542578	2	C	T	NA	*ABCC2*	10	0.207726808	0.161407065	0.311177117	0.039921571	0.303023983	0.377740239	0.30865503	0.046128273	0.3153561	0.2548448
34	rs7910642	101541579	2	G	A	NA	*ABCC2*	10	0.107563025	0.973268313	0.085613019	0.139654975	0.086548488	0.653323267	0.085198267	0.140225746	0.08975572	0.1931096
35	rs4148386	101548468	3	A	G	NA	*ABCC2*	10	0.442324855	0.129130664	0.197533946	0.127175191	0.195307612	0.748719341	0.195566267	0.235417397	0.1982107	0.02506933
36	rs8187707	101610533	3	C	T	NA	*ABCC2*	10	0.058770556	0.095507915	0.111571486	0.07259935	0.087486966	0.684987702	0.103644469	0.367928484	0.1087784	*< 1*10-8*
37	rs1051640	48768486	2	A	G	NA	*ABCC3*	17	0.174780527	0.852401095	0.146415735	0.645960596	0.143899896	0.056472335	0.144166766	0.029936152	0.1448096	0.2241819
38	rs4148416	48753423	2	C	T	NA	*ABCC3*	17	0.060174968	0.940128442	0.441094249	0.186294547	0.439311783	0.99394779	0.442781323	0.196792225	0.4416392	0.1517074
39	rs4793665	48712087	2	T	C	NA	*ABCC3*	17	0.420598743	0.36962251	0.117811502	0.208097343	0.11282586	0.951319054	0.116681521	0.283380615	0.1177531	0.02568038
40	rs9895420	48712038	3	T	A	NA	*ABCC3*	17	0.090893231	0.266780672	0.286142173	0.247470752	0.28540146	0.822052068	0.284395473	0.264360884	0.288644	0.0245887
41	rs17268122	95844494	2	G	T	NA	*ABCC4*	13	0.223215283	0.55827223	0.13296226	0.636842513	0.133159541	0.48501968	0.132879274	0.492250561	0.1121012	0.07763722
42	rs17268282	95925314	2	G	T	NA	*ABCC4*	13	0.073071341	0.250930324	0.199880192	0.661813515	0.201876955	0.464304497	0.199229442	0.765589725	0.2088848	0.3073711
43	rs2274407	95859035	2	C	A	NA	*ABCC4*	13	0.064939837	0.099298759	0.076477636	0.348483565	0.077163712	0.068417009	0.077983847	0.065293998	0.07126842	0.9478854
44	rs4773866	95929326	2	C	T	NA	*ABCC4*	13	0.085462901	0.852078426	0.197758586	0.418684193	0.191553702	0.368869572	0.197067399	0.606488974	0.2006881	0.8779732
45	rs7317112	95923523	2	A	G	NA	*ABCC4*	13	0.26405499	0.980723634	0.205121805	0.881875446	0.20354536	0.66104501	0.205734729	0.429715107	0.2049396	0.06990433
46	rs868853	95955076	2	T	C	NA	*ABCC4*	13	0.076552882	0.532197411	0.107727636	0.269430207	0.110427529	0.234991542	0.108101687	0.732843725	0.1081523	0.6268031
47	rs899494	95861804	2	G	A	NA	*ABCC4*	13	0.145206399	0.123306562	0.173672125	0.046882079	0.172888425	0.486105404	0.17389164	0.253294714	0.1778635	0.4193322
48	rs9556455	95697416	2	G	A	NA	*ABCC4*	13	0.138228197	0.86770208	0.060927516	0.695680548	0.05620438	0.648311284	0.060130894	0.986077558	0.05424158	0.3159807
49	rs9561778	95713715	2	G	T	NA	*ABCC4*	13	0.21940999	0.282914366	0.419304113	0.373373191	0.432429614	0.724888401	0.421631573	0.187364123	0.4358378	0.9398552
50	rs2274405	95858978	3	C	T	NA	*ABCC4*	13	0.352284151	0.186668956	0.219698482	0.696026038	0.224504692	0.077730382	0.221259422	0.500748828	0.2233558	0.3216945
51	rs3742106	95673791	3	A	C	NA	*ABCC4*	13	0.402181208	0.906654535	0.07430611	0.146401961	0.075703858	0.533543229	0.074219677	0.045790725	0.07158207	0.2736217
52	rs9516519	95672457	3	T	G	NA	*ABCC4*	13	0.144551354	0.619655985	0.064247204	0.356586872	0.066840459	0.35989113	0.064685413	0.449838682	0.06602204	0.1649555
53	rs3805111	183689989	2	G	A	NA	*ABCC5*	3	0.063707426	0.971490999	0.084464856	0.486029018	0.079979145	0.122675688	0.084098132	0.217885996	0.08645253	0.9139518
54	rs8180093	183642303	2	G	A	NA	*ABCC5*	3	0.061990394	0.699205033	0.26936901	0.637786824	0.278519291	0.966970989	0.269365466	0.611301331	0.2743199	0.543975
55	rs10937158	183708439	3	C	T	NA	*ABCC5*	3	0.474935804	0.284712426	0.08206869	0.953391641	0.075495308	0.066630466	0.079641945	0.813279943	0.07843594	0.5459893
56	rs2293001	183700211	3	C	T	NA	*ABCC5*	3	0.347772554	0.147099256	0.142097644	0.600218174	0.146715328	0.537177681	0.143620495	0.514180448	0.1282055	0.08194285
57	rs3749442	183660585	3	G	A	NA	*ABCC5*	3	0.167206885	0.720507134	0.129642572	0.742862951	0.128467153	0.754695505	0.131704036	0.85655181	0.1338329	0.957173
58	rs3792581	183695870	3	C	A	NA	*ABCC5*	3	0.176726542	0.912600688	0.226837061	0.06194851	0.223044838	0.988130418	0.224180912	0.048284908	0.2191852	0.4976166
59	rs2231142	89052323	1	G	T	NA	*ABCG2*	4	0.121948348	0.861449106	0.064047524	0.971045716	0.068821689	0.852493133	0.064506557	0.906908681	0.06606492	0.6538555
60	rs17731538	89055379	2	G	A	NA	*ABCG2*	4	0.156237748	0.932176553	0.060927516	0.721020586	0.059541189	0.362908852	0.061083086	0.685132434	0.0646645	0.956648
61	rs2199939	39911494	2	C	T	NA	*ABCG2*	8	0.170652913	0.625413351	0.15235623	0.058556677	0.153701773	0.447351159	0.153853333	0.285797648	0.1671135	0.05905998
62	rs2231135	89079994	2	A	G	NA	*ABCG2*	4	0.055699045	0.815900339	0.162265375	0.957815882	0.16016684	0.781973304	0.164184218	0.677084712	0.1649755	0.8861808
63	rs2231148	89028478	2	T	A	NA	*ABCG2*	4	0.364139841	0.551347093	0.056384784	0.10482826	0.056934307	0.983030317	0.056281945	0.147154809	0.05713432	0.5932062
64	rs2725264	89026109	2	T	C	NA	*ABCG2*	4	0.060995774	0.093339184	0.360872604	0.821513952	0.368300313	0.746892559	0.362856886	0.739104704	0.3712469	0.3293736
65	rs7699188	89096061	2	G	A	NA	*ABCG2*	4	0.150295792	0.728743652	0.063248802	0.26262389	0.062877998	0.528303855	0.062512278	0.043978442	0.06158675	0.2160949
66	rs17731799	89068455	3	G	T	NA	*ABCG2*	4	0.479218153	0.641297461	0.143320687	0.486667414	0.152137643	0.397984176	0.146413102	0.694408402	0.1530467	0.6143459
67	rs2622604	89078924	3	C	T	NA	*ABCG2*	4	0.256063485	0.869677019	0.269219249	0.843446206	0.287799791	0.035975729	0.27102593	0.314538202	0.2757951	0.9799706
68	rs2725256	89050998	3	A	G	NA	*ABCG2*	4	0.383376217	0.051900661	0.194988019	0.894094242	0.189468196	0.206097655	0.194503307	0.820371154	0.2067039	0.01577387
69	rs4148157	89020934	3	G	A	NA	*ABCG2*	4	0.102210255	0.938677137	0.32435603	0.030500873	0.329092805	0.010151105	0.325760649	0.015913366	0.3225237	0.1726202
70	rs12443580	55866337	2	T	C	NA	*CES1*	16	0.26405208	0.809500229	0.055061901	0.109282461	0.062565172	0.241139977	0.056488527	0.098109196	0.05050589	0.02524929
71	rs3815583	55867042	2	A	C	NA	*CES1*	16	0.196593218	0.433577703	0.44513778	0.690535157	0.438164755	0.858391666	0.444461599	0.880480961	0.4485127	0.7292141
72	rs2244613	55844609	3	T	G	NA	*CES1*	16	0.189174107	0.968591568	0.185952476	0.321295193	0.196037539	0.723859859	0.187598248	0.207030135	0.185798	0.9430286
73	rs2244614	55844617	3	A	G	NA	*CES1*	16	0.400883838	0.917584771	0.422299321	0.155338835	0.420020855	0.984620397	0.423595607	0.587456235	0.4220394	0.002322937
74	rs2606345	75017176	2	A	C	NA	*CYP1A1*	15	0.326463853	0.841730261	0.201203075	0.140484163	0.206256517	0.306414988	0.203633259	0.543934763	0.2063229	0.01893043
75	rs2472297	75027880	3	C	T	NA	*CYP1A1*	15	0.277331443	0.961463521	0.088857827	0.578028944	0.09092805	0.49588861	0.089627572	0.521588815	0.08490263	0.1149129
76	rs4646903	75011641	3	A	G	NA	*CYP1A1*	15	0.102604401	0.234710407	0.447683706	0.1388589	0.442961418	0.387222336	0.444398229	0.029979333	0.4393946	0.4203743
77	rs762551	75041917	1	A	C	*1F	*CYP1A2*	15	0.258330067	0.059398042	0.365864617	0.970831094	0.373722628	0.209742982	0.368456236	0.906245777	0.374327	0.1281515
78	rs11636419	75047600	2	A	G	NA	*CYP1A2*	15	0.056704507	0.821731843	0.052840455	0.101693901	0.052033368	0.651210168	0.05137131	0.02160369	0.05761093	0.7937895
79	rs2472299	75033400	3	G	A	NA	*CYP1A2*	15	0.261323085	0.060128845	0.066843051	0.023247851	0.075078206	0.046722317	0.06795165	0.000152791	0.06801547	*< 1*10-8*
80	rs1056836	38298203	2	C	G	NA	*CYP1B1*	2	0.446308725	0.876055112	0.2657498	0.974380337	0.271845673	0.383288369	0.26707057	0.511971005	0.2611868	0.09634747
81	rs1800440	38298139	2	T	C	NA	*CYP1B1*	2	0.186954123	0.625017274	0.050344449	0.352092845	0.05318	0.7889	0.05056493	0.352739239	0.0565631	0.472139739
82	rs10012	38302390	3	G	C	NA	*CYP1B1*	2	0.273656575	0.712800255	0.129442891	0.792143121	0.134306569	0.16983355	0.129348148	0.801336959	0.1336075	0.2320686
83	rs28399433	41356379	1	A	C	*9A/*9B/*13/*15	*CYP2A6*	19	0.063116824	0.020251284	0.09377496	0.577530639	0.098957247	0.674730048	0.09406882	0.811003189	0.1007263	0.1893324
84	rs8192726	41354496	1	C	A	*9b	*CYP2A6*	19	0.062029217	0.047022858	0.069064497	0.184232504	0.068821689	0.852493133	0.068304315	0.511604728	0.06958865	0.1777295
85	rs56113850	41353107	2	C	T	NA	*CYP2A6*	19	0.429158233	0.251934724	0.19521266	0.929721144	0.202711157	0.393661999	0.197456389	0.951314891	0.2088051	0.00839187
86	rs8192725	41354712	3	G	A	*1B	*CYP2A6*	19	0.247542857	0.144649505	0.122928315	0.801100018	0.116996872	0.126998361	0.118759298	0.020307186	0.1214999	0.7818508
87	rs2279343	41515263	1	A	G	*4	*CYP2B6*	19	0.245391842	0.770206525	0.097144569	0.772525626	0.093847758	0.003284807	0.096950241	0.210490209	0.0941616	0.2769282
88	rs1042389	41524153	2	T	C	NA	*CYP2B6*	19	0.208183818	0.939143594	0.393245807	0.660646886	0.403441084	0.62420847	0.394083613	0.916160112	0.3951876	0.8570128
89	rs2279342	41510127	2	A	T	NA	*CYP2B6*	19	0.090602106	0.870423847	0.217302316	0.413899678	0.221584984	0.356150439	0.216033521	0.457989899	0.2354793	0.1675238
90	rs2054675	41495755	3	T	C	NA	*CYP2B6*	19	0.240561253	0.722355894	0.393645168	0.230238603	0.389051095	0.4887597	0.392100336	0.133281622	0.4063358	0.1101212
91	rs2279344	41515483	3	A	G	NA	*CYP2B6*	19	0.401451232	0.182526369	0.238942692	0.147380104	0.238373306	0.337727023	0.238517002	0.055272553	0.2415883	0.2749666
92	rs4803419	41512792	3	C	T	NA	*CYP2B6*	19	0.29914966	0.947185691	0.39858726	0.062744329	0.407194995	0.87300222	0.399965303	0.473470324	0.4015661	0.8360004
93	rs7254579	41494891	3	T	C	NA	*CYP2B6*	19	0.335237455	0.412772839	0.285493211	0.4084298	0.304588113	0.384485251	0.289351852	0.042903124	0.2930131	0.9883079
94	rs12248560	96521657	1	C	T	*17	*CYP2C19*	10	0.211574953	0.721646969	0.072484026	0.279648298	0.079979145	0.081878646	0.07358412	0.242728155	0.07693872	0.7701147
95	rs4244285	96541616	1	G	A	*2	*CYP2C19*	10	0.139307649	0.388993244	0.059804313	0.321906435	0.055787278	0.003648692	0.059994369	0.593012415	0.06345643	0.7760909
96	rs3814637	96521045	3	C	T	NA	*CYP2C19*	10	0.06532411	0.959318786	0.340854633	0.795103481	0.338894682	0.656065496	0.337302923	0.651698545	0.3377004	0.8104226
97	rs7916649	96534584	3	G	A	NA	*CYP2C19*	10	0.416228904	0.304800189	0.38266274	0.397032658	0.381751825	0.286609752	0.383478287	0.857800009	0.3710704	0.1530958
98	rs10509681	96798749	1	T	C	*3	*CYP2C8*	10	0.121223219	0.07165618	0.412739617	0.639047691	0.417413973	0.257282675	0.411331512	0.895110578	0.4046622	0.5538406
99	rs1058930	96818119	1	G	C	*4	*CYP2C8*	10	0.051251647	0.541177061	0.313623203	0.370752487	0.310531804	0.490354382	0.312334906	0.100506337	0.3188762	0.4150226
100	rs11572080	96827030	1	C	T	NA	*CYP2C8*	10	0.12246839	0.054787706	0.403878794	0.813477123	0.400625652	0.462293884	0.402021889	0.666633824	0.4082129	0.8496473
101	rs17110453	96829529	1	A	C	*1C	*CYP2C8*	10	0.124761477	0.082735361	0.09771865	0.287965976	0.096767466	0.35102783	0.096724097	0.085822523	0.09529492	0.3393734
102	rs7909236	96829430	1	G	T	*1B	*CYP2C8*	10	0.249302377	0.595223987	0.184754393	0.801750909	0.184567258	0.636955207	0.183663175	0.677403741	0.1888402	0.3611373
103	rs193451	108310986	2	A	G	NA	*CYP2C8*	7	0.438117145	0.168587074	0.146066294	0.031568704	0.146402503	0.068079356	0.142403051	0.343312188	0.1508659	2.56E-06
104	rs1113129	96811045	3	G	C	NA	*CYP2C8*	10	0.205330204	0.077090815	0.462060703	0.981852079	0.468717414	0.516651537	0.465157673	0.511991234	0.4874342	0.9759373
105	rs1057910	96741053	1	A	C	*3	*CYP2C9*	10	0.065329535	0.670182216	0.127570887	0.028290836	0.135140772	0.018972266	0.130658558	0.016353069	0.1318848	0.1656866
106	rs1799853	96702047	1	C	T	*2	*CYP2C9*	10	0.129086913	0.036191833	0.248627196	0.4751529	0.252346194	0.52740754	0.248507174	0.282398584	0.2322301	0.333007
107	rs4918758	96697252	2	T	C	NA	*CYP2C9*	10	0.370946022	0.603029874	0.288213858	0.195328151	0.288112617	0.415051974	0.28919328	0.095778716	0.2956853	0.3056005
108	rs9332197	96740908	2	T	C	NA	*CYP2C9*	10	0.047484011	0.156213	0.201852037	0.636701067	0.207612096	0.911540371	0.203522737	0.810814376	0.1992852	0.1556136
109	rs7089580	96705223	3	A	T	NA	*CYP2C9*	10	0.208836014	0.98971349	0.05486222	0.26195174	0.055891554	0.900029323	0.055341976	0.417971189	0.05406515	0.2786757
110	rs1065852	42526694	1	G	A	*10	*CYP2D6*	22	0.222339058	0.179421402	0.215679912	0.715067908	0.214702815	0.972864547	0.214779705	0.625579889	0.2257315	0.9968395
111	rs1135840	42522613	1	G	C	NA	*CYP2D6*	22	0.444718021	0.003907004	0.058107029	0.0892396	0.054535975	0.61641805	0.056933262	0.485329745	0.05339373	0.9977362
112	rs16947	42523943	1	G	A	*2	*CYP2D6*	22	0.32235023	0.000135292	0.449680511	0.22061876	0.447758081	0.414873087	0.449147001	0.75578823	0.4594282	0.5488002
113	rs28371725	42523805	1	C	T	*41	*CYP2D6*	22	0.09677968	0.397679961	0.173597244	0.378338526	0.170072993	0.521455135	0.171878709	0.322018499	0.1741842	0.1095684
114	rs3892097	42524947	1	C	T	*4	*CYP2D6*	22	0.207740368	0.178521614	0.437025759	0.362929348	0.437643379	0.058610191	0.438417727	0.535236964	0.4370312	0.7240854
115	rs1080983	42528568	3	C	T	NA	*CYP2D6*	22	0.321003352	0.000114991	0.057832468	0.647219432	0.062252346	0.97288391	0.058720047	0.185716996	0.05617497	0.5887881
116	rs1080985	42528382	3	C	G	NA	*CYP2D6*	22	0.221194961	0.051390526	0.095896565	0.475247247	0.103023983	0.702744696	0.095606538	0.882243595	0.1015472	0.4065006
117	rs8192766	135339385	2	T	G	NA	*CYP2E1*	10	0.066450022	0.017402258	0.349640575	0.720921109	0.351094891	0.671831655	0.350637982	0.768538585	0.3297941	0.1894587
118	rs2242480	99361466	1	C	T	*1G	*CYP3A4*	7	0.082084507	0.002193865	0.404328075	0.131022901	0.403023983	0.929716368	0.403582985	0.200671934	0.4058335	0.06744483
119	rs2246709	99365719	2	A	G	NA	*CYP3A4*	7	0.267556549	0.448833833	0.142072684	0.056508269	0.142127216	0.427304389	0.142763841	0.37055315	0.1425589	0.8156045
120	rs35599367	99366316	2	G	A	NA	*CYP3A4*	7	0.046788991	0.522293144	0.458566294	0.109539932	0.468925965	0.994212691	0.464484928	0.056436574	0.4465487	0.3475424
121	rs3735451	99355975	3	T	C	NA	*CYP3A4*	7	0.09637746	0.006376792	0.359999002	0.764973817	0.351824818	0.844956893	0.3555417	0.536334013	0.3722907	0.5819374
122	rs2257401	99306685	3	G	C	NA	*CYP3A7*	7	0.079437688	0.060867437	0.173397564	0.79478724	0.17080292	0.943711631	0.171336143	0.894707275	0.1723831	0.7975453
123	rs10306135	125137695	2	A	T	NA	*PTGS1*	9	0.127213458	0.49243238	0.181634385	0.302371944	0.179353493	0.983164805	0.179961436	0.420258453	0.1815367	0.0329312
124	rs12353214	125161870	2	C	T	NA	*PTGS1*	9	0.092031492	0.942952427	0.486621406	0.56759361	0.485922836	0.251740909	0.484650361	0.244659698	0.4682546	0.1597321
125	rs1236913	125133479	2	C	T	NA	*PTGS1*	9	0.066791394	0.774778531	0.251996805	0.206866123	0.248383733	0.6194861	0.2525611	0.2869569	0.2553499	0.2678104
126	rs1330344	125131688	2	T	C	NA	*PTGS1*	9	0.200013492	0.711145774	0.387604832	0.588475086	0.391240876	0.923106327	0.387104435	0.138788537	0.380708	0.2836147
127	rs5788	125143792	2	C	A	NA	*PTGS1*	9	0.110233555	0.003638534	0.108950679	0.840489904	0.099687174	0.530774337	0.105867499	0.674692807	0.09868632	0.1250888
128	rs3842787	125133507	3	C	T	NA	*PTGS1*	9	0.064431541	0.072131869	0.192566893	0.280461009	0.193743483	0.739810106	0.191601191	0.339197442	0.1773967	0.1006764
129	rs12208357	160543148	1	C	T	NA	*SLC22A1*	6	0.065971145	0.507203817	0.368235823	0.054451131	0.384254432	0.598356606	0.377866667	0.055405938	0.3762641	0.08530785
130	rs461473	160543562	2	G	A	NA	*SLC22A1*	6	0.097641099	0.473126673	0.277031749	0.145988056	0.284567258	0.149677542	0.278089362	0.122489528	0.2669111	0.1748926
131	rs622342	160572866	2	A	C	NA	*SLC22A1*	6	0.391997805	0.572424176	0.099740415	0.858510955	0.108446298	0.986257469	0.101723369	0.641872626	0.09975471	0.08244419
132	rs683369	160551204	2	C	G	NA	*SLC22A1*	6	0.210473684	0.507204141	0.259409944	0.77927162	0.252763295	0.433385923	0.259193856	0.316053619	0.2720386	0.6275928
133	rs628031	160560845	3	G	A	NA	*SLC22A1*	6	0.396829682	0.238705364	0.275284545	0.816719645	0.277163712	0.250225308	0.275238887	0.65345275	0.2772396	0.2230025
134	rs11231809	64302950	2	T	A	NA	*SLC22A11*	11	0.389775281	0.103918889	0.249375998	0.378094399	0.248905109	0.296108901	0.248768546	0.669364526	0.2572772	0.000666108
135	rs316003	160645832	2	T	C	NA	*SLC22A2*	6	0.237688442	0.465599156	0.235647963	0.557140022	0.243378519	0.901762196	0.238146936	0.825955764	0.239993	0.08311653
136	rs316019	160670282	3	C	A	NA	*SLC22A2*	6	0.103078615	0.195243945	0.415809704	0.638849444	0.395933264	0.252767103	0.409204424	0.922150258	0.4093866	0.8276689
137	rs2076828	160872786	2	C	G	NA	*SLC22A3*	6	0.399253145	0.28348065	0.290684904	0.785406361	0.303649635	0.260002658	0.294753776	0.457274567	0.2947823	0.2819627
138	rs555754	160769423	3	G	A	NA	*SLC22A3*	6	0.467919856	0.127463691	0.330770767	0.567255071	0.339207508	0.140253221	0.333125464	0.143682023	0.3252902	0.7645817
139	rs1050152	131676320	3	C	T	NA	*SLC22A4*	5	0.440009265	0.879877745	0.0626248	0.720956582	0.064233577	0.019198525	0.063616138	0.197164888	0.06325918	0.3767132
140	rs2631372	131703578	2	G	C	NA	*SLC22A5*	5	0.289641301	0.081985351	0.426093251	0.344708696	0.432533889	0.56824179	0.424229833	0.28757554	0.4277281	0.1504987
141	rs4149170	62752289	2	C	T	NA	*SLC22A6*	11	0.072735487	0.490464652	0.208915735	0.243680546	0.214911366	0.923329153	0.208808167	0.067687679	0.2139033	0.3589543
142	rs11568563	21457434	2	T	G	NA	*SLCO1A2*	12	0.062378168	0.026745087	0.213658147	0.862153546	0.211783107	0.403819045	0.211949909	0.936609169	0.2115962	0.9985906
143	rs10841795	21487544	3	A	G	NA	*SLCO1A2*	12	0.121358331	0.268796518	0.327600839	0.045028646	0.330031283	0.827400641	0.326172226	0.628157751	0.3246384	*< 1*10-8*
144	rs4148978	21488675	3	C	T	NA	*SLCO1A2*	12	0.315465533	0.197434782	0.226287939	0.346022696	0.226798749	0.991341866	0.22507485	0.836259157	0.2280778	*< 1*10-8*
145	rs4148981	21487068	3	C	T	NA	*SLCO1A2*	12	0.318014089	0.169501462	0.093350639	0.560257274	0.100625652	0.202093265	0.095247989	0.013940384	0.104979	0.6654468
146	rs2306283	21329738	1	A	G	*1B	*SLCO1B1*	12	0.381736252	0.463774366	0.077276358	0.75414184	0.084045881	0.103919807	0.078834174	0.153399731	0.08264701	0.966784
147	rs4149056	21331549	1	T	C	*5	*SLCO1B1*	12	0.158961319	0.06127064	0.065944489	0.279982701	0.067674661	0.909082284	0.065758645	0.078475131	0.07035441	0.2179228
148	rs4149032	21317791	2	C	T	NA	*SLCO1B1*	12	0.329846775	0.985200024	0.39434405	0.068866719	0.402189781	0.426340069	0.396078067	0.704767191	0.4154233	0.05026729
149	rs4149057	21331599	2	C	T	*18/*19	*SLCO1B1*	12	0.385931559	0.409869136	0.10153754	0.243049599	0.104588113	0.123125668	0.102366731	0.023750365	0.1034855	0.707992
150	rs4149087	21392562	2	T	G	NA	*SLCO1B1*	12	0.404636577	0.703323462	0.471046326	0.524969679	0.469760167	0.885080787	0.470057388	0.22417665	0.4668787	0.2891047
151	rs11045818	21329761	3	G	A	*18	*SLCO1B1*	12	0.1467106	0.359287815	0.438698083	0.607482328	0.4111	0.09884	0.435682766	0.527501128	0.4318939	0.6622558
152	rs11045872	21372344	3	A	G	NA	*SLCO1B1*	12	0.171932567	0.046742552	0.123127995	0.400063009	0.122523462	0.729605688	0.122501965	0.140661612	0.1314976	0.3878525
153	rs11045879	21382619	3	T	C	NA	*SLCO1B1*	12	0.172230907	0.176456412	0.325404353	0.83522154	0.319499479	0.461657387	0.321779533	0.549002299	0.3194753	0.1498299
154	rs34671512	21391976	3	A	C	*19/*20/*22/*35	*SLCO1B1*	12	0.052112934	0.42100996	0.327900359	0.852536019	0.322732013	0.689928031	0.32438409	0.447160605	0.3218547	0.1278852
155	rs3829306	21292280	3	C	T	NA	*SLCO1B1*	12	0.053631285	0.024728974	0.15218151	0.493984504	0.154014599	0.971346915	0.150979228	0.94072363	0.1614003	0.08323085
156	rs4149036	21327740	3	C	A	NA	*SLCO1B1*	12	0.19875672	0.494746734	0.178239816	0.803088302	0.178415016	0.62644209	0.17662592	0.153369025	0.1865074	0.4444142
157	rs4149118	21011581	2	A	G	NA	*SLCO1B3*	12	0.310250391	0.190322009	0.172124601	0.135768317	0.172471324	0.981097097	0.172269531	0.663904763	0.160436	0.2827952
158	rs4149117	21011480	3	G	T	NA	*SLCO1B3*	12	0.142053253	0.990872327	0.0545627	0.32653818	0.051407716	0.960979971	0.053371102	0.245016465	0.05200801	0.4780846
159	rs7977213	20980800	3	C	G	NA	*SLCO1B3*	12	0.451160433	0.754442221	0.058206869	0.637211762	0.055161627	0.794740312	0.056592575	0.111333608	0.05003541	0.9422934
160	rs3794271	20860093	2	A	G	NA	*SLCO1C1*	12	0.398702481	0.3067926	0.203649161	0.015067249	0.197184567	0.705484476	0.201469933	0.098974476	0.1937534	0.4129888
161	rs12422149	74883577	2	G	A	NA	*SLCO2B1*	11	0.095031729	0.313938256	0.140949481	0.01786053	0.138686131	0.686796695	0.140935082	0.044552316	0.1413373	0.7041944
162	rs2712807	74861881	2	A	G	NA	*SLCO2B1*	11	0.180605181	0.903006823	0.455621006	0.276881852	0.440250261	0.587539099	0.452249948	0.381982334	0.4543974	0.6944677

Appendices (continue):

Supplementary material 3: The 1st prize award certificate.

SCHOOL OF MEDICINE

POSTGRADUATE RESEARCH STUDENT SYMPOSIUM 2019

1st PRIZE AWARD

MUSTAFA MALKI

Posters Category 2

Presented by: Prof John Hayes
Date: 7th June 2019

University of Dundee

This page was left intentionally

293

281

This page was left intentionally